(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 October 2017 (19.10.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/180319 A1

(51) International Patent Classification:
C12Q 1/68 (2006.01) *C40B 20/00* (2006.01)

(21) International Application Number:
PCT/US2017/024314

(22) International Filing Date:
27 March 2017 (27.03.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/322,343 14 April 2016 (14.04.2016) US

(71) Applicant: **QUANTAPORE, INC.** [US/US]; 1455 Adams Drive, Suite 1081, Menlo Park, California 94025 (US).

(72) Inventors: **ANDERSON, Brett N.**; 1455 Adams Drive, Suite 1081, Menlo Park, California 94025 (US). **HUBER, Martin**; 1455 Adams Drive, Suite 1081, Menlo Park, California 94025 (US). **MACEVICZ, Stephen C.**; 1455 Adams Drive, Suite 1081, Menlo Park, California 94025 (US).

(74) Agents: **LEVINE, David A.** et al.; Levine Bagade Han LLP, 2400 Geng Rd, Suite 120, Palo Alto, California 94303 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

— *with international search report (Art. 21(3))*

(54) **Title**: MIXED OPTICAL SIGNALS IN POLYMER ANALYSIS WITH NANOPORES

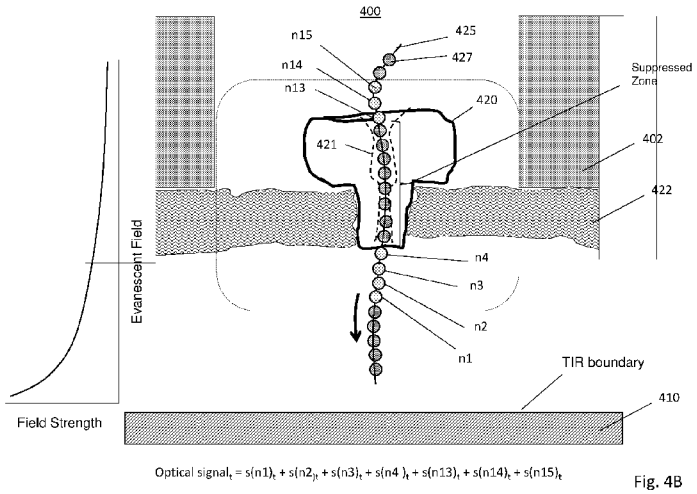


Fig. 4B

(57) **Abstract**: The invention is directed to nanopore-based methods for analyzing polymers, such as polynucleotides or proteins, containing optical labels specific for different kinds of monomers. In some embodiments, methods of the invention include steps of (a) translocating a polymer through a nanopore, wherein different kinds of monomers of the polymer are labeled with different optical labels that generate distinguishable optical signals and wherein the nanopore constrains the monomers to move single file through an excitation zone that encompasses a plurality of monomers; (b) detecting a time-ordered set of optical signals from the monomers as the polymer passes through the excitation zone; (c) separating optical signals from different kinds of monomers to form monomer-specific time-ordered sets of optical signals; and (d) determining a sequence of monomers from the monomer-specific time-ordered sets of optical signals from the polymer.

WO 2017/180319 A1

MIXED OPTICAL SIGNALS IN POLYMER ANALYSIS WITH NANOPORES

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit of priority to U.S. Provisional Patent Application No. 62/322,343, filed on April 14, 2016, the content of which is incorporated herein by reference in its entirety.

BACKGROUND

[0002] Nanopore sequencing has been proposed as an approach to overcome a host of challenges in current DNA sequencing technologies, including reduction of per-run sequencing cost, simplification of sample preparation, reduction of run times, increasing sequence read lengths, providing real-time sample analysis, and the like. However, polymer analysis, such as DNA analysis, with nanopores has its own set of technical difficulties, such as, reliable nanostructure fabrication, control of DNA translocation rates, unambiguous nucleotide discrimination, detection and processing of signals from large arrays of nanoscale sensors, and so on, e.g. Branton et al, Nature Biotechnology, 26(10): 1146-1153 (2008).

[0003] Optical detection of nucleotides has been proposed as a potential solution to some of the technical difficulties in the field of nanopore sequencing, for example, the difficulty of collecting independent signals from large arrays of nanopores. However, with fluorescence-based signals, overcoming background noise in the optical detection of single molecules remains a significant challenge, which has led to the frequent use of microscopy systems, such as total internal reflection fluorescence (TIRF) systems, which minimize background excitation by limiting the spatial region of excitation. However, even with currently available techniques for limiting excitation volume, collected signals at any instant may comprise contributions from multiple optical labels within the same resolution limited area and excitation volume, which greatly complicates base calling.

[0004] In view of the above, it would be advantageous to nanopore sequencing technology and its particular applications, such as optically-based nanopore sequencing, if methods and devices were available that would permit unambiguous base-calling despite detected optical signals comprising light from multiple spatially indistinguishable labels.

SUMMARY OF THE INVENTION

[0005] The present invention is directed to methods and devices for polymer analysis, especially polynucleotide analysis, using optical labels and nanopores.

[0006] In some embodiments, the invention is directed to a method of analyzing a polymer comprising the steps of (a) translocating a polymer through a nanopore, wherein different kinds of monomers of the polymer are labeled with different optical labels that generate distinguishable optical signals and wherein the nanopore constrains the monomers to move single file through an excitation zone that encompasses a plurality of monomers; (b) detecting a time-ordered set of optical signals from the monomers as the polymer passes through the excitation zone; (c) separating optical signals from different kinds of monomers to form monomer-specific time-ordered sets of optical signals; and (d) determining a sequence of monomers from the monomer-specific time-ordered sets of optical signals from the polymer.

[0007] In other embodiments, the invention is directed to a method of analyzing a polynucleotide comprising the steps of (a) translocating a polynucleotide through a nanopore, nucleotides of the polynucleotide being labeled with fluorescent labels and the nanopore having a bore that spatially constrains the fluorescent labels to prevent emission of fluorescent signals during translocation thereof; (b) exciting the fluorescent labels; (c) detecting a time series of fluorescent signals from the fluorescent labels as the polynucleotide translocates through the bore; and (d) determining a sequence of fluorescent labels attached to nucleotides of the polynucleotide from the time series of fluorescent signals.

[0008] The present invention advantageously overcomes the problem of optical measurements containing contributions of more than one optical label in optically-based nanopore analysis. These and other advantages of the present invention are exemplified in a number of implementations and applications, some of which are summarized below and throughout the specification.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] Figs. 1 illustrates elements of the invention in an embodiment using FRET and an epi-illumination system.

[0010] Fig. 2A illustrates elements of the invention in an embodiment using FRET and a TIRF system.

[0011] Fig. 2B illustrates elements of the invention in an embodiment wherein two labeled monomers emit optical signals at the same time as they pass through an excitation zone.

[0012] Fig. 3 illustrates the basic components of a confocal epi-illumination system.

[0013] Fig. 4A-4B illustrate elements of a TIRF system for excitation without FRET.

[0014] Fig. 5 is a flow chart illustrating a step for calling nucleotide sequences based on measurements of optical signals comprising light from multiple optical labels.

[0015] Figs. 6A-6C illustrate embodiments of the invention employing quenching agents in a trans chamber, a cis chamber and in both cis and trans chambers, respectively.

[0016] Fig. 7 illustrates an embodiment of the invention using a protein nanopore and epi-illumination with a metal layer on the nanopore array to reduce background or TIR with FRET.

DETAILED DESCRIPTION OF THE INVENTION

[0017] While the invention is amenable to various modifications and alternative forms, specifics thereof have been shown by way of example in the drawings and will be described in detail. It should be understood, however, that the intention is not to limit the invention to the particular embodiments described. On the contrary, the intention is to cover all modifications, equivalents, and alternatives falling within the spirit and scope of the invention. For example, particular nanopore types and numbers, particular labels, FRET pairs, detection schemes, fabrication approaches of the invention are shown for purposes of illustration. It should be appreciated, however, that the disclosure is not intended to be limiting in this respect, as other types of nanopores, arrays of nanopores, and other fabrication technologies may be utilized to implement various aspects of the systems discussed herein. Guidance for aspects of the invention is found in many available references and treatises well known to those with ordinary skill in the art, including, for example, Cao, Nanostructures & Nanomaterials (Imperial College Press, 2004); Levinson, Principles of Lithography, Second Edition (SPIE Press, 2005); Doering and Nishi, Editors, Handbook of Semiconductor Manufacturing Technology, Second Edition (CRC Press, 2007); Sawyer et al, Electrochemistry for Chemists, 2nd edition (Wiley Interscience, 1995); Bard and Faulkner, Electrochemical Methods: Fundamentals and Applications, 2nd edition (Wiley, 2000); Lakowicz, Principles of Fluorescence Spectroscopy, 3rd edition (Springer, 2006); Hermanson, Bioconjugate Techniques, Second Edition (Academic Press, 2008); and the like, which relevant parts are hereby incorporated by reference.

[0018] In one aspect, the invention is directed to methods and devices for analyzing polymers using nanopores and optical detection. In some embodiments, different kinds of monomer have different labels that generate distinguishable optical signals which allow identification of monomers. Such labeled polymers may be translocated through nanopores that constrain the monomers to move single file through an optical detection region in, or intersecting, a resolution limited area. A series of optical signals is measured from such resolution limited areas wherein each optical measurement comprises a plurality of component signals from different adjacent monomers (whose order in the polymer cannot be determined from single measurements because, for example, the component signals are generated from within a diffraction limited area). In part, the invention is based on a recognition and appreciation that in some configurations, optically-based nanopore analysis of polymers (i) generates a time series

of optical measurements that comprise overlapping contributions from sequences of more than one labeled monomer, thereby making it difficult, if not impossible, to determine an ordering of the monomers from a single measurement, and (ii) by selecting optical labels for monomers which generate distinguishable signals, the optical measurements can be separated into contributions from different labels on different kinds of monomers, which allows overlapping measurements to be converted into sequence information.

[0019] In one aspect, a method of the invention may be implemented by the following steps: (a) translocating a polymer through a nanopore, wherein different kinds of monomers of the polymer are labeled with different optical labels that generate distinguishable optical signals and wherein the nanopore constrains the monomers to move single file through an excitation zone or signal generation zone that encompasses a plurality of monomers; (b) detecting a time series of optical signals from the monomers as the polymer passes through the excitation zone or signal generation zone; (c) separating optical signals from different kinds of monomers; and (d) determining a sequence of monomers from time series of the separated optical signals from the excitation zone or the signal generation zone.

[0020] As used herein, the terms "excitation zone," "signal generation zone," "detection zone," or like terms mean a spatial region or volume adjacent to a nanopore from which optical signals are generated or collected. In some embodiments, a labeled monomer may transit such zones during a "transition interval". Such regions or volumes may be determined by several factors known to those of skill in the art, including, but not limited to, the manner in which excitation energy is delivered to optical labels (e.g. FRET, TIRF, or the like), the nature of the optical labels (e.g. whether there is self- and/or mutual quenching), the nanopore configuration employed (e.g. protein or non-protein nanopores, presence of absence of light-blocking layer), presence or absence of quenching agents, and so on. In some embodiments, excitation or signal generation zones may be determined empirically by measuring optical signals generated by test or calibration polynucleotides which have known sequences of nucleotides, known labels and/or known concentrations of quenching agents. In some embodiments, excitation zones or signal generation zones adjacent to nanopores each has a volume and geometry so that a number of nucleotides occupying such volume is in the range of from 1 to 4, or in the range of from 1 to 3, or in still other embodiments, in the range of from 1 to 2. In some embodiments, such zones comprise a single contiguous volume adjacent to the trans opening of a nanopore.

[0021] In some embodiments, the invention relates to the use of nanopores, fluorescent quenching, and fluorescent signaling to sequentially identify nucleotides of fluorescently labeled polynucleotide analytes. Such analysis of polynucleotide analytes may be carried out on single polynucleotides or on pluralities of polynucleotides in parallel at the same time, for example, by using an array of nanopores. In some embodiments, nucleotides are labeled with fluorescent labels that are capable of at least three

states while attached to a polynucleotide: (i) A substantially quenched state wherein fluorescence of an attached fluorescent label is quenched by a fluorescent label on an immediately adjacent monomer or by interaction with a quenching agent; for example, a fluorescent label attached to a polynucleotide in accordance with the invention is substantially quenched when the labeled polynucleotide is free in conventional aqueous solutions or buffers for studying and manipulating the polynucleotide. (ii) A sterically constrained state while a labeled polynucleotide is translocating through a nanopore such that the free-solution movements or alignments of attached fluorescent labels are disrupted or limited so that there is little or no detectable fluorescent signal generated from the fluorescent label. (iii) A transition state wherein fluorescent labels attached to a polynucleotide transition from the sterically constrained state to a quenched state as the nucleotide of the fluorescent label exits the nanopore (during a "transition interval" or "interval"). Some embodiments of the invention are (in part) applications of the discovery that during the transition interval a fluorescent label (on an otherwise substantially fully labeled and self-quenched or quenched polynucleotide) is capable of generating a detectable fluorescent signal and that the number of exiting labels contributing to a measured signal may be (at least in part) controlled by controlling the translocation speed of the labeled polynucleotide. If translocation speed (e.g. nucleotides exiting a nanopore per msec) is higher than the transition rate (from signal-capable to quenched, i.e. the quenching rate), then measured fluorescent signals, or signal samples, may contain contributions from more than one label. This circumstance makes signal analysis more difficult and possibly less accurate. In accordance with some embodiments, this problem may be ameliorated by adjusting translocation speed, for example by reducing translocation speed, so that substantially only one fluorescent label at a time contributes fluorescence to a measured fluorescent signal.

[0022] Without the intention of being limited by any theory underlying this discovery, it is believed that the fluorescent signal generated during the transition interval is due to the presence of one or more freely rotatable dipoles of the fluorescent labels that emerged from a nanopore, which renders the fluorescent labels capable of generating a fluorescent signal, for example, after direct excitation or via excitation via FRET. In some embodiments, the polynucleotide is a single stranded polynucleotide, such as, DNA or RNA, but especially a single stranded DNA. In some embodiments, the invention includes a method for determining a nucleotide sequence of a polynucleotide by recording signals generated by fluorescent labels as they exit a nanopore one at a time as a polynucleotide translocates through the nanopore. In some embodiments, a translocation speed may be selected to minimize the number of fluorescent labels that contribute to measured fluorescent signals. Such selection may be made either by real-time adjustment of parameters controllable during operation (such as the voltage across the nanopores, temperature, or the like) or by predetermined instrument set-up (e.g. reaction

buffer viscosity, ion concentration, or the like). Upon exit, each attached fluorescent label transitions during a transition interval from a constrained state in the nanopore to a quenched state on the polynucleotide in free solution. During the transition interval the label is capable of generating a fluorescent signal which can be measured. In other words, in some embodiments, a step of the method of the invention comprises exciting each fluorescent label as it is transitioning from a constrained state in the nanopore to a quenched state on the polymer in free solution. As mentioned above, during this transition interval or period a fluorescent label is capable of emitting a detectable fluorescent signal indicative of the nucleotide to which it is attached.

[0023] In some embodiments, "substantially quenched" as used above means a fluorescent label generates a fluorescent signal at least thirty percent reduced from a signal generated under the same conditions, but without adjacent mutually quenching labels. In some embodiments, "substantially quenched" as used above means a fluorescent label generates a fluorescent signal at least fifty percent reduced from a signal generated under the same conditions, but without adjacent mutually quenching labels.

[0024] In some embodiments, optical signals may be FRET signals or they may be fluorescent emissions from directly excited fluorescent labels attached to monomers. Fig. 1 illustrates components of one embodiment in which a protein nanopore (100) is disposed in a lipid bilayer (102) disposed (in turn) across aperture (104) of solid state membrane (106), which comprises opaque layer (108) (such as a metal layer), silicon nitride layer (110) and silicon support layer (112). Opaque layer (108) prevents or reduces transmission of excitation beam (114) through solid state membrane (106) where it could excite undesired background fluorescence. As polymer (120) with differently labeled monomers (illustrated as black (122) and white (124)) pass through nanopore (100), at each measurement interval a plurality of monomers (such as, 141, 142 and 143) are present in excitation zones (126) and (128) within the same resolution limited area. In the illustrated embodiment, optical measurements are made with an epi-illumination system and it is assumed that nanopore (100) has been selected so that optical signals from monomers interior to nanopore (100) are suppressed and do not contribute to measured optical signals. Excitation zone (128) is a FRET zone adjacent to FRET donor (130); that is, excitation zone (128) defines a distance from FRET donor (130) within which FRET can occur between FRET donor (130) and an optical label attached to a monomer, which may also be referred to as an acceptor label, or FRET acceptor label. Excitation zone (126) is a non-propagating protrusion of a component of excitation beam (114) into aperture (104) which occurs whenever the dimensions of aperture (104) are selected to be sufficiently below the wavelength of excitation beam (114). As illustrated, in this embodiment, a plurality of monomers (141, 142 and 143) would contribute to an

optical signal measured at the instant, or interval, during which monomers (141, 142 and 143) are in the excitation zones (126) and (128).

[0025] Fig. 2 illustrates an embodiment in which optical measurements are made with total internal reflection fluorescence (TIRF) excitation in a system such as described in Soni et al, Review of Scientific Instruments, 81: 014301 (2010); and in U.S. patent publication 2012/0135410, which is incorporated herein by reference. In this embodiment, protein nanopore (200) with attached FRET donor (202) is inserted into lipid bilayer (204) disposed on solid state membrane (206) with aperture (208). Total internal reflection (TIR) is made possible by selecting electrolytes on cis (205) and trans (207) sides of solid state membrane (206) with different indices of refraction. As a result, TIR boundary (210) is created at or near the plane that solid state membrane (206) is disposed in, so that an evanescent field is created on the cis (205) side of solid state membrane (206). The evanescent field may excite optical labels prior to their entry into nanopore (200). FRET donor (202) is excited directly by light reflected at the TIR boundary (210), so that FRET can take place between FRET donor (202) and labels on monomers (219) within FRET zone (220). As in the embodiment of Fig. 1, nanopore (200) may be selected so that fluorescent emissions by labels are suppressed when labeled monomers are in the bore of nanopore (200). A plurality of monomers, such as 225, 226 and 227, contribute to an optical measurement recorded at the indicated configuration in the figure.

[0026] In some embodiments, labels on monomers may be excited by an evanescence field alone using an apparatus similar to that shown in Fig. 4A. In this apparatus, a very narrow second chamber on the trans side of a nanopore or nanopore array permits an evanescent field to extend from a surface of an underlying glass slide to establish excitation zones both at entrances and exits of the nanopores, so that each optical measurement associated with a nanopore contains contributions from a plurality of labeled monomers. Array of apertures (400) (which may include protein nanopores inserted in a lipid bilayer), may be formed in silicon nitride layer (402), which may have a thickness in the range of from 20-100 nm. Silicon nitride layer (402) may be formed on a silicon support layer (403). Second chamber (406) may be formed by silicon nitride layer (402), silicon dioxide layer (404) which determines the height of second chamber (406), and surface (408) of glass slide (410). Silicon dioxide layer (404) may have a thickness in the range of from 50-100 nm. A desired evanescent field (407) extending from surface (408) across silicon nitride layer (402) may be established by directing light beam (412) at an appropriate angle relative to glass slide (410) so that TIR occurs. For driving labeled polynucleotide analytes through array (400), cis(-) conditions may be established in first chamber (416) and trans(+) conditions may be established in second chamber (406) with electrodes operationally connected to first and second chambers (406 and 421). Fig. 4B is a close-up of a particular embodiment of an aperture in array (400) which diagrammatically shows protein nanopore (420)

inserted in lipid bilayer (422) that is disposed on a surface of silicon nitride layer (402). Polymer (425) with labeled monomers (for example, 427) is shown translocating through nanopore (420), which has been selected with bore (421) having dimensions that cause suppression of fluorescent emissions of labels interior to bore (421). In this embodiment, a measured optical signal at a particular time point, t, or interval, from a resolution limited area containing aperture (400) may comprise contributions from a plurality of labels on monomers in the excitation regions, for example, monomers n_1-n_4 and n_{13}-n_{15}.

[0027] Fig. 2B provides a further illustration of collecting fluorescent signals that comprise fluorescence generated by two labeled monomers. Fig. 2B shows labeled polynucleotide (2000) translocating through nanopore (2002), wherein labeled polynucleotide (2000) comprises two labels "a" and "b" (for example, which may correspond to dC being labeled with "a" and dA, dG and dT being labeled with "b", or the like). Labels of nucleotides free of nanopore (2002) are quenched, either by interaction with other labels (2011) or by action of quenching agents (not shown). Labels of nucleotides inside of nanopore (2002) are constrained and/or oriented (2014) so that they produce no detectable signal during all or part of their transit through the nanopore. As nucleotides of labeled polynucleotide (2000) emerge from exit (2015) of nanopore (2002) they become capable of being excited by excitation beam (2010) and generating a detectable signal for an interval prior to being quenched. If translocation speed V_1 is high then the distance (2008) traveled by a nucleotide prior to quenching may exceed the inter-nucleotide distance of polynucleotide (2000) so that more than one label contributes fluorescence to a fluorescent signal collected by detector (2018), i.e. a measured fluorescent signal. Since the distance between adjacent labels is below the diffraction limit of excitation light (2010) no information is obtained about the ordering of the labels (in the excitation zone or signal generation zone (if defined by quenching) (2099)), although there are approaches to deduce such information using specialized algorithms, e.g. Timp et al, Biophys. J., 102: L37-L39 (2012); Carson et al, Nanotechnology, 26: 074004 (2015). In the case of optical detection using fluorescent labels with distinct emission bands, measured fluorescent signals may be separated into two or more channels, e.g. using bandpass filters, in order to assess the relative contributions of fluorescence from multiple labels. However, as the number of fluorescent labels contributing fluorescence increases, e.g. 3, 4, or more, the difficulty in determining a correct ordering of nucleotides increases. The signal intensities for two channels, e.g. corresponding to emission maxima of two fluorescent labels, is illustrated in Fig. 2B (2031 and 2032) where two fluorescent labels contribute to a measured signal. Intensity values represented by solid lines, e.g. 2033, are from label "a," and intensity values represented by dashed lines, e.g. 2036, are from label "b". The presence of solid and dashed lines in both channels of Fig. 2B reflects overlapping emission bands of the fluorescent labels, which

when collected together complicates analysis because amounts of a measured intensity are from both labels.

Sequence Determination

[0028] In accordance with the invention, when a labeled polymer translocates through a nanopore and its associated excitation zones, a time-ordered set of optical measurements are recorded. Optical measurements at adjacent time points are overlapping in the sense that each optical measurement contains contributions from labels of adjacent monomers. Thus, for example, if three monomers generate signals at each time point (for example, B, C and D of polymer …-A-(B-C-D)-… moving through an excitation zone from left to right), and if one monomer exits the excitation zone and another monomer enters the excitation zone (indicated by parentheses) between successive measurements (for example, A enters and D exits: -(A-B-C)-D…), then two successive optical measurements will contain contributions from the same monomers (in this example, both measurements include contributions from B and C. The above example is based on a very simplified model of polymer translocation through nanopores; however, the concept of successive overlapping optical measurements is applicable to more complex descriptions of polymer translocation.

[0029] Since emissions from a plurality of different labeled monomers at a nanopore originate from the same resolution limited area, relative position information (in particular, sequence information) about the monomers cannot be determined from a single optical measurement. However, because of the overlap and the use of labels that generate monomer-specific signals, in some embodiments, sequence information may be determined from the time-ordered set of optical signal measurements when it is separated into a plurality of time-ordered sets of monomer-specific signals. Algorithms similar to those used in sequencing-by-hybridization (SBH) to reconstruct target polynucleotide sequences from hybridization data may be used to reconstruct target polynucleotides here, e.g. U.S. patent 5,002,867; Timp et al, Biophys. J., 102: L37-L39 (2012); or the like, which are incorporated by reference. The constraints of (i) time-ordered overlapping signals and signals and (ii) their separation into monomer-specific components significantly simplify the determination step in the case of optical detection.

[0030] Fig. 5 illustrates one embodiment of a step for determining monomer sequence information from a time-ordered set of overlapping optical signals based on a simple model of nanopore translocation. The simple model assumes that optical measurements at each time step (except at the entry and exit of a polymer from a nanopore) each contain signal contributions from the same number of monomers (referred to in Fig. 5 as an "n-tuple" to indicate that a measurement would contain contributions from n monomers). It is understood that more complex models may allow for differing

numbers of contributing monomers in each measurement, for local variations in translocation speed, deviations in linear movement of monomers, and other like phenomena. That is, in some embodiments, optical measurements at different times may have contributions from different numbers of nucleotides. In some embodiments, the differing number of nucleotides are ordered along a segment of the target polynucleotide. The step of determining illustrated by Fig. 5 assumes that a labeled polymer has passed through a nanopore and that a time ordered set of optical measurements has been made, including separation of optical signals into monomer-specific signals (500). The entry and exit of a polymer are treated differently since there are necessarily different numbers of monomers in the excitation zone(s) upon entry and exit. In this embodiment, it is assumed that initial and final optical measurements under these conditions permits the initial and final monomers to be determined directly from their monomer-specific signal. In other embodiments, preparation of labeled polymers for analysis may include insertion of a plurality of predetermined labeled nucleotides at one or both ends of such labeled polymers for the purpose of generating a known sequence of optical signals to aid in a sequence determination step. Such predetermined labeled nucleotides would be similar to key sequences in Ion Torrent or 454 sequencing, e.g. U.S. patent 7,575,865, which is incorporated by reference.

[0031] Returning to Fig. 5, at the beginning of a determining step, time index, i, is set to zero; the index, j, for candidate sequences at the current time, i, is set to 1 (502); and the initial n-tuple of the set of monomer-specific time-ordered optical signals is examined (504). Such examination comprises first determining from the measurement at time i all possible n-tuples of monomers that are consistent with the measurement, then determining from those n-tuples which ones that properly overlap candidate sequence Si. New candidate sequences Si+1 are formed (and a sequence Si is extended) by each properly overlapping n-tuple for the set consistent with the measurement (506). New extended candidate sequences, Si+1, are stored and the index giving the number of candidate sequences at time i+1, Ji+1, is updated (508). This step is repeated until every candidate sequence, Si, has been examined (510), and a similar examination is carried out at each time, i, until each optical measurement in the time-ordered set has been examined.

Nanopores and Nanopore Arrays

[0032] Nanopores used with the invention may be solid-state nanopores, protein nanopores, or hybrid nanopores comprising protein nanopores or organic nanotubes such as carbon or graphene nanotubes, configured in a solid-state membrane, or like framework. Important features of nanopores include constraining polymer analytes, such as polynucleotides, so that their monomers pass through a signal generation region (or excitation zone, or the like) in sequence, That is, so that monomers, such as

nucleotides, pass through a detection zone (or excitation region or like region) in single file. In some embodiments, additional features of nanopores include passing single stranded nucleic acids while not passing double stranded nucleic acids, or equivalently bulky molecules. In other embodiments, nanopores, especially protein nanopores, may be selected so that their bores are sized so that labels of monomers are sterically constrained so that FRET signals, or even fluorescent signals, are suppressed.

[0033] In some embodiments, nanopores used in connection with the methods and devices of the invention are provided in the form of arrays, such as an array of clusters of nanopores, which may be disposed regularly on a planar surface. In some embodiments, clusters are each in a separate resolution limited area so that optical signals from nanopores of different clusters are distinguishable by the optical detection system employed, but optical signals from nanopores within the same cluster cannot necessarily be assigned to a specific nanopore within such cluster by the optical detection system employed.

[0034] Solid state nanopores may be fabricated in a variety of materials including but not limited to, silicon nitride (Si_3N_4), silicon dioxide (SiO_2), and the like. The fabrication and operation of nanopores for analytical applications, such as DNA sequencing, are disclosed in the following exemplary references that are incorporated by reference: Ling, U.S. patent 7,678,562; Hu et al, U.S. patent 7,397,232; Golovchenko et al, U.S. patent 6,464,842; Chu et al, U.S. patent 5,798,042; Sauer et al, U.S. patent 7,001,792; Su et al, U.S. patent 7,744,816; Church et al, U.S. patent 5,795,782; Bayley et al, U.S. patent 6,426,231; Akeson et al, U.S. patent 7,189,503; Bayley et al, U.S. patent 6,916,665; Akeson et al, U.S. patent 6,267,872; Meller et al, U.S. patent publication 2009/0029477; Howorka et al, International patent publication WO2009/007743; Brown et al, International patent publication WO2011/067559; Meller et al, International patent publication WO2009/020682; Polonsky et al, International patent publication WO2008/092760; Van der Zaag et al, International patent publication WO2010/007537; Yan et al, Nano Letters, 5(6): 1129-1134 (2005); Iqbal et al, Nature Nanotechnology, 2: 243-248 (2007); Wanunu et al, Nano Letters, 7(6): 1580-1585 (2007); Dekker, Nature Nanotechnology, 2: 209-215 (2007); Storm et al, Nature Materials, 2: 537-540 (2003); Wu et al, Electrophoresis, 29(13): 2754-2759 (2008); Nakane et al, Electrophoresis, 23: 2592-2601 (2002); Zhe et al, J. Micromech. Microeng., 17: 304-313 (2007); Henriquez et al, The Analyst, 129: 478-482 (2004); Jagtiani et al, J. Micromech. Microeng., 16: 1530-1539 (2006); Nakane et al, J. Phys. Condens. Matter, 15 R1365-R1393 (2003); DeBlois et al, Rev. Sci. Instruments, 41(7): 909-916 (1970); Clarke et al, Nature Nanotechnology, 4(4): 265-270 (2009); Bayley et al, U.S. patent publication 2003/0215881; and the like.

[0035] In some embodiments, the invention comprises nanopore arrays with one or more light-blocking layers, that is, one or more opaque layers. Typically nanopore arrays are fabricated in thin

sheets of material, such as, silicon, silicon nitride, silicon oxide, aluminum oxide, or the like, which readily transmit light, particularly at the thicknesses used, e.g. less than 50-100 nm. For electrical detection of analytes this is not a problem. However, in optically-based detection of labeled molecules translocating nanopores, light transmitted through an array invariably excites materials outside of intended reaction sites, thus generates optical noise, for example, from nonspecific background fluorescence, fluorescence from labels of molecules that have not yet entered a nanopore, or the like. In one aspect, the invention addresses this problem by providing nanopore arrays with one or more light-blocking layers that reflect and/or absorb light from an excitation beam, thereby reducing background noise for optical signals generated at intended reaction sites associated with nanopores of an array. In some embodiments, this permits optical labels in intended reaction sites to be excited by direct illumination. In some embodiments, an opaque layer may be a metal layer. Such metal layer may comprise Sn, Al, V, Ti, Ni, Mo, Ta, W, Au, Ag or Cu. In some embodiments such metal layer may comprise Al, Au, Ag or Cu. In still other embodiments, such metal layer may comprise aluminum or gold, or may comprise solely aluminum. The thickness of an opaque layer may vary widely and depends on the physical and chemical properties of material composing the layer. In some embodiments, the thickness of an opaque layer may be at least 5 nm, or at least 10 nm, or at least 40 nm. In other embodiments, the thickness of an opaque layer may be in the range of from 5-100 nm; in other embodiments, the thickness of an opaque layer may be in the range of from 10-80 nm. An opaque layer need not block (i.e. reflect or absorb) 100 percent of the light from an excitation beam. In some embodiments, an opaque layer may block at least 10 percent of incident light from an excitation beam; in other embodiments, an opaque layer may block at least 50 percent of incident light from an excitation beam.

[0036] Opaque layers or coatings may be fabricated on solid state membranes by a variety of techniques known in the art. Material deposition techniques may be used including chemical vapor deposition, electrodeposition, epitaxy, thermal oxidation, physical vapor deposition, including evaporation and sputtering, casting, and the like. In some embodiments, atomic layer deposition may be used, e.g. U.S. patent 6,464,842; Wei et al, Small, 6(13): 1406-1414 (2010), which are incorporated by reference.

[0037] In some embodiments, a 1-100 nm channel or aperture may be formed through a solid substrate, usually a planar substrate, such as a membrane, through which an analyte, such as single stranded DNA, is induced to translocate. In other embodiments, a 2-50 nm channel or aperture is formed through a substrate; and in still other embodiments, a 2-30 nm, or a 2-20 nm, or a 3-30 nm, or a 3-20 nm, or a 3-10 nm channel or aperture if formed through a substrate. The solid-state approach of generating nanopores offers robustness and durability as well as the ability to tune the size and shape of

the nanopore, the ability to fabricate high-density arrays of nanopores on a wafer scale, superior mechanical, chemical and thermal characteristics compared with lipid-based systems, and the possibility of integrating with electronic or optical readout techniques. Biological nanopores on the other hand provide reproducible narrow bores, or lumens, especially in the 1-10 nanometer range, as well as techniques for tailoring the physical and/or chemical properties of the nanopore and for directly or indirectly attaching groups or elements, such as fluorescent labels, which may be FRET donors or acceptors, by conventional protein engineering methods. Protein nanopores typically rely on delicate lipid bilayers for mechanical support, and the fabrication of solid-state nanopores with precise dimensions remains challenging. In some embodiments, solid-state nanopores may be combined with a biological nanopore to form a so-called "hybrid" nanopore that overcomes some of these shortcomings, thereby providing the precision of a biological pore protein with the stability of a solid state nanopore. For optical read out techniques a hybrid nanopore provides a precise location of the nanopore which simplifies the data acquisition greatly.

[0038] In some embodiments, clusters may also be formed by disposing protein nanopores in lipid bilayers supported by solid phase membrane containing an array of apertures. For example, such an array may comprise apertures fabricated (e.g. drilled, etched, or the like) in solid phase support. The geometry of such apertures may vary depending on the fabrication techniques employed. In some embodiments, each such aperture is associated with, or encompassed by, a separate resolution limited area; however, in other embodiments, multiple apertures may be within the same resolution limited area. The cross-sectional area of the apertures may vary widely and may or may not be the same as between different clusters, although such areas are usually substantially the same as a result of conventional fabrication approaches. In some embodiments, apertures have a minimal linear dimension (e.g. diameter in the case of circular apertures) in the range of from 10 to 200 nm, or have areas in the range of from about 100 to $3x10^4$ nm^2. Across the apertures may be disposed a lipid bilayer. The distribution of protein nanopores per aperture may be varied, for example, by controlling the concentration of protein nanopores during inserting step. In such embodiments, clusters of nanopores may comprise a random number of nanopores. In some embodiments, in which protein nanopores insert randomly into apertures, clusters containing one or more apertures on average have a number of protein nanopores that is greater than zero; in other embodiments, such clusters have a number of protein nanopores that is greater than 0.25; in other embodiments, such clusters have a number of protein nanopores that is greater than 0.5; in other embodiments, such clusters have a number of protein nanopores that is greater than 0.75; in other embodiments, such clusters have a number of protein nanopores that is greater than 1.0.

[0039] In some embodiments, methods and devices of the invention comprise a solid phase membrane, such as a SiN membrane, having an array of apertures therethrough providing communication between a first chamber and a second chamber (also sometimes referred to as a "cis chamber" and a "trans chamber") and supporting a lipid bilayer on a surface facing the second, or trans, chamber. In some embodiments, diameters of the aperture in such a solid phase membrane may be in the range of 10 to 200 nm, or in the range of 20 to 100 nm. In some embodiments, such solid phase membranes further include protein nanopores inserted into the lipid bilayer in regions where such bilayer spans the apertures on the surface facing the trans chamber. In some embodiments, such protein nanopores are inserted from the cis side of the solid phase membrane using techniques described herein. In some embodiments, such protein nanopores have a structure identical to, or similar to, α-hemolysin in that it comprises a barrel, or bore, along an axis and at one end has a "cap" structure and at the other end has a "stem" structure (using the terminology from Song et al, Science, 274: 1859-1866 (1996)). In some embodiments using such protein nanopores, insertion into the lipid bilayer results in the protein nanopore being oriented so that its cap structure is exposed to the cis chamber and its stem structure is exposed to the trans chamber.

[0040] In some embodiments, the present invention may employ hybrid nanopores in clusters, particularly for optical-based nanopore sequencing of polynucleotides. Such nanopores comprise a solid-state orifice, or aperture, into which a protein biosensor, such as a protein nanopore, is stably inserted. A charged polymer may be attached to a protein nanopore (e.g. alpha hemolysin) by conventional protein engineering techniques after which an applied electric field may be used to guide a protein nanopore into an aperture in a solid-state membrane. In some embodiments, the aperture in the solid-state substrate is selected to be slightly smaller than the protein, thereby preventing it from translocating through the aperture. Instead, the protein will be embedded into the solid-state orifice.

[0041] In some embodiments, a donor fluorophore is attached to the protein nanopore. This complex is then inserted into a solid-state aperture or nanohole (for example, 3-10 nm in diameter) by applying an electric field across the solid state nanohole, or aperture, until the protein nanopore is transported into the solid-state nanohole to form a hybrid nanopore. The formation of the hybrid nanopore can be verified by (a) the inserted protein nanopore causing a drop in current based on a partial blockage of the solid-state nanohole and by (b) the optical detection of the donor fluorophore.

[0042] Solid state, or synthetic, nanopores may be preprared in a variety of ways, as exemplified in the references cited above. In some embodiments a helium ion microscope may be used to drill the synthetic nanopores in a variety of materials, e.g. as disclosed by Yang et al, Nanotechnolgy, 22: 285310 (2011), which is incorporated herein by reference. A chip that supports one or more regions of a thin-film material, e.g. silicon nitride, that has been processed to be a free-standing membrane is

introduced to the helium ion microscope (HIM) chamber. HIM motor controls are used to bring a free-standing membrane into the path of the ion beam while the microscope is set for low magnification. Beam parameters including focus and stigmation are adjusted at a region adjacent to the free-standing membrane, but on the solid substrate. Once the parameters have been properly fixed, the chip position is moved such that the free-standing membrane region is centered on the ion beam scan region and the beam is blanked. The HIM field of view is set to a dimension (in μm) that is sufficient to contain the entire anticipated nanopore pattern and sufficient to be useful in future optical readout (i.e. dependent on optical magnification, camera resolution, etc.). The ion beam is then rastered once through the entire field of view at a pixel dwell time that results in a total ion dose sufficient to remove all or most of the membrane autofluorescence. The field of view is then set to the proper value (smaller than that used above) to perform lithographically-defined milling of either a single nanopore or an array of nanopores. The pixel dwell time of the pattern is set to result in nanopores of one or more predetermined diameters, determined through the use of a calibration sample prior to sample processing. This entire process is repeated for each desired region on a single chip and/or for each chip introduced into the HIM chamber.

[0043] In some embodiments, a nanopore may have one or more labels attached for use in optically-based nanopore sequencing methods. The label may be a member of a Forster Resonance Energy Transfer (FRET) pair. Such labels may comprise organic fluorophores, chemiluminescent labels, quantum dots, metallic nanoparticles and/or fluorescent proteins. Target nucleic acids may have one distinct label per nucleotide. The labels attached to the nucleotides may be selected from the group consisting of organic fluorophores. The label attachment site in the pore protein can be generated by conventional protein engineering methods, e.g. a mutant protein can be constructed that will allow the specific binding of the label. As an example, a cysteine residue may be inserted at the desired position of the protein which inserts a thiol (SH) group that can be used to attach a label. The cysteine can either replace a natural occurring amino acid or can be incorporated as an addition amino acid. A maleimide-activated label is then covalently attached to the thiol residue of the protein nanopore. In a preferred embodiment the attachment of the label to the protein nanopore or the label on the nucleic acid is reversible. By implementing a cleavable crosslinker, an easily breakable chemical bond (e.g. an S-S bond or a pH labile bond) is introduced and the label may be removed when the corresponding conditions are met.

[0044] In some embodiments, an epi-illumination system, in which excitation beam delivery and optical signal collection occurs through a single objective, may be used for direct illumination of labels on a polymer analyte or donors on nanopores. The basic components of a confocal epi-illumination system for use with the invention is illustrated in Fig. 3. Excitation beam (302) passes through dichroic

(304) and onto objective lens (306) which focuses (310) excitation beam (302) onto layered membrane (300), in which labels are excited directly to emit an optical signal, such as a fluorescent signal, of are excited indirectly via a FRET interaction to emit an optical signal. Such optical signal is collected by objective lens (306) and directed to dichroic (304), which is selected so that it passes light of excitation beam (302) but reflects light of optical signals (311). Reflected optical signals (311) passes through lens (314) which focuses it through pinhole (316) and onto detector (318).

[0045] In some embodiments, a device for implementing the above methods for analyzing polymers (such as single stranded polynucleotides) typically includes a set of electrodes for establishing an electric field across the layered membrane and nanopores. Single stranded nucleic acids are exposed to nanopores by placing them in an electrolyte in a first chamber, which is configured as the "cis" side of the layered membrane by placement of a negative electrode in the chamber. Upon application of an electric field, the negatively single stranded nucleic acids are captured by nanopores and translocated to a second chamber on the other side of the layered membrane, which is configured as the "trans" side of membrane by placement of a positive electrode in the chamber. The speed of translocation depends in part on the ionic strength of the electrolytes in the first and second chambers and the applied voltage across the nanopores. In optically based detection, a translocation speed may be selected by preliminary calibration measurements, for example, using predetermined standards of labeled single stranded nucleic acids that generate signals at different expected rates per nanopore for different voltages. Thus, for DNA sequencing applications, a translocation speed may be selected based on the signal rates from such calibration measurements. Consequently, from such measurements a voltage may be selected that permits, or maximizes, reliable nucleotide identifications, for example, over an array of nanopores. In some embodiments, such calibrations may be made using nucleic acids from the sample of templates being analyzed (instead of, or in addition to, predetermined standard sequences). In some embodiments, such calibrations may be carried out in real time during a sequencing run and the applied voltage may be modified in real time based on such measurements, for example, to maximize the acquisition of nucleotide-specific signals.

Controlling Translocation Speed

[0046] The role of translocation speed of polynucleotides through nanopores and the need for its control have been appreciated in the field of nanopore technology wherein changes in electric current are use to identify translocating analytes. A wide variety of methods have been used to control translocation speed, which include both methods that can be adjusted in real-time without significant difficulty (e.g. voltage potential across nanopores, temperature, and the like) and methods that can be adjusted during operation only with difficulty (reaction buffer viscosity, presence or absence of charged

side chains in the bore of a protein nanopore, ionic composition and concentration of the reaction buffer, velocity-retarding groups attached or hybridized to polynucleotide analytes, molecular motors, and the like), e.g. Bates et al, Biophysical J., 84: 2366-2372 (2003); Carson et al, Nanotechnology, 26(7): 074004 (2015); Yeh et al, Electrophoresis, 33(23): 58-65 (2012); Meller, J. Phys. Cond. Matter, 15: R581-R607 (2003); Luan et al, Nanoscale, 4(4): 1068-1077 (2012); Keyser, J. R. Soc. Interface, 8: 1369-1378 (2011); and the like, which are incorporated herein by reference. In some embodiments, a step or steps are included for active control of translocation speed while a method of the invention is being implemented, e.g. voltage potential, temperature, or the like; in other embodiments, a step or steps are included that determine a translocation speed that is not actively controlled or changed while a method of the invention is being implemented, e.g. reaction buffer viscosity, ionic concentration, and the like. In regard to the latter, in some embodiments, a translocation speed is selected by providing a reaction buffer having a concentration of glycerol, or equivalent reagent, in the range of from 1 to 60 percent.

[0047] In regard to the former embodiments (with real-time translocation speed adjustment), a measure of whether one or more than one label is contributing fluorescence to measured signals may be based on the distribution of fluorescence intensity among a plurality of channels over which fluorescence is collected. Typically the plurality of channels include 2, 3, or 4 channels corresponding to the emission bands of the fluorescent labels used. In a measured sample of fluorescence emanating from a region adjacent to a nanopore exit, if only a single label contributes to a measured signal, the relative distribution of signal intensity among the different channels (e.g. 4 channels) could be represented ideally as (1,0,0,0); (0,1,0,0); (0,0,1,0) or (0,0,0,1). On the other hand, if more than one label contributed to a measured fluorescent signal, the relative distributions would include non-zero values in more than one channel, with a worse case being four different labels contributing equally, which would appear as (.25,.25,.25,.25) in the above representation. A measure which would vary monotonically between a maximum value corresponding to relative intensity distributions (1,0,0,0); (0,1,0,0); (0,0,1,0) or (0,0,0,1) and a minimum value corresponding to a relative intensity distribution of (.25,.25,.25,.25) may be used for controlling in real-time a translocation speed. For example, an initial translocation speed could be lowered based on the value of such a measure that was near its minimum. Such lowering may be implemented, for example, by lowering a potential voltage across the nanopores by a predetermined amount, after which the measure could be re-calculated. Such steps could be repeated until the process was optimized.

[0048] As mentioned above, translocation speeds depend in part on the voltage difference (or electrical field strength) across a nanopore and conditions in the reaction mixture, or buffer, of a first chamber where polynucleotides are exposed to the nanopores (e.g. disposed in a solid phase membrane

making up one wall of the first chamber). Polynucleotide capture rates by nanopores depend on concentration of such polynucleotides. In some embodiments, conventional reaction mixture conditions for nanopore sequencing may be employed with the invention (for controlling translocatin speed by varying voltage potential across nanopores), for example, 1M KCl (or equivalent salt, such as NaCl, LiCl, or the like) and a pH buffering system (which, for example, ensures that proteins being used, e.g. protein nanopores, nucleases, or the like, are not denatured). In some embodiments, a pH buffering system may be used to keep the pH substantially constant at a value in the range of 6.8 to 8.8. In some embodiments, a voltage difference across the nanopores may be in the range of from 70 to 200 mV. In other embodiments, a voltage difference across the nanopores may be in the range of from 80 to 150 mV. An appropriate voltage for operation may be selected using conventional measurement techniques. Current (or voltage) across a nanopore may readily be measured using commercially available instruments. A voltage difference may be selected so that translocation speed is within a desired range. In some embodiments, a range of translocation speeds comprises those speeds less than 1000 nucleotides per second. In other embodiments, a range of translocation speeds is from 10 to 800 nucleotides per second; in other embodiments, a range of translocation speeds is from 10 to 600 nucleotides per second; in other embodiments, a range of translocation speeds is from 200 to 800 nucleotides per second; in other embodiments, a range of translocation speeds is from 200 to 500 nucleotides per second. Likewise, other factors affecting translocation speed, e.g. temperature, viscosity, ion concentration, charged side chains in the bore of a protein nanopore, and the like, may be selected to obtain translocation speeds in the ranges cited above.

[0049] In some embodiments, a device for implementing the above methods for single stranded nucleic acids typically includes providing a set of electrodes for establishing an electric field across the nanopores (which may comprise an array). Single stranded nucleic acids are exposed to nanopores by placing them in an electrolyte (i.e. reaction buffer) in a first chamber, which is configured as the "cis" side of the layered membrane by placement of a negative electrode in the chamber. Upon application of an electric field, the negatively single stranded nucleic acids are captured by nanopores and translocated to a second chamber on the other side of the layered membrane, which is configured as the "trans" side of membrane by placement of a positive electrode in the chamber. As mentioned above, the speed of translocation depends in part on the ionic strength of the electrolytes in the first and second chambers and the applied voltage across the nanopores. In optically based detection, a translocation speed may be selected by preliminary calibration measurements, for example, using predetermined standards of labeled single stranded nucleic acids that generate signals at different expected rates per nanopore for different voltages. Thus, for DNA sequencing applications, an initial translocation speed may be selected based on the signal rates from such calibration measurements, as well as the measure

based on relative signal intensity distribution discussed above. Consequently, from such measurements a voltage may be selected that permits, or maximizes, reliable nucleotide identifications, for example, over an array of nanopores. In some embodiments, such calibrations may be made using nucleic acids from the sample of templates being analyzed (instead of, or in addition to, predetermined standard sequences). In some embodiments, such calibrations may be carried out in real time during a sequencing run and the applied voltage may be modified in real time based on such measurements, for example, to maximize the acquisition of nucleotide-specific signals.

Embodiments Employing Mutually and Self-Quenching Labels

[0050] As mentioned above, in some embodiments, self- and mutually quenching fluorescent labels may be used in addition to quenching agents in order to reduce fluorescent emissions outside of those from labels on nucleotides exiting nanopores, i.e. to restrict the spatial extent of a signal generation zone. Use of such fluorescent labels is disclosed in U.S. patent publication 2016/0122812, which is incorporated by reference. In some embodiments, monomers are labeled with fluorescent labels that are capable of at least three states while attached to a target polynucleotide: (i) A substantially quenched state wherein fluorescence of an attached fluorescent label is quenched by a fluorescent label on an immediately adjacent monomer; for example, a fluorescent label attached to a polynucleotide in accordance with the invention is substantially quenched when the labeled polynucleotide is free in conventional aqueous solution for studying and manipulating the polynucleotide. (ii) A sterically constrained state wherein a labeled polynucleotide is translocating through a nanopore such that the free-solution movements or alignments of an attached fluorescent label is disrupted or limited so that there is little or no detectable fluorescent signal generated from the fluorescent label. (iii) A transition state wherein a fluorescent label attached to a polynucleotide transitions from the sterically constrained state to the quenched state as the fluorescent label exits the nanopore (during a "transition interval") while the polynucleotide translocates through the nanopore.

[0051] In part, this example is an application of the discovery that during the transition interval a fluorescent label (on an otherwise substantially fully labeled and self-quenched polynucleotide) is capable of generating a detectable fluorescent signal. Without the intention of being limited by any theory underlying this discovery, it is believed that the fluorescent signal generated during the transition interval is due to the presence of a freely rotatable dipole in the fluorescent label emerging from the nanopore, which renders the fluorescent label temporarily capable of generating a fluorescent signal, for example, after direct excitation or via FRET. In both the sterically constrained state as well as the quenched state, the dipoles are limited in their rotational freedom thereby reducing or limiting the number of emitted photons. In some embodiments, the polynucleotide is a polynucleotide, usually

a single stranded polynucleotide, such as, DNA or RNA, but especially single stranded DNA. In some embodiments, the invention includes a method for determining a nucleotide sequence of a polynucleotide by recording signals generated by attached fluorescent labels as they exit a nanopore one at a time as a polynucleotide translocates through the nanopore. Upon exit, each attached fluorescent label transitions during a transition interval from a constrained state in the nanopore to a quenched state on the polynucleotide in free solution. In other words, in some embodiments, a step of the method of the invention comprises exciting each fluorescent label as it is transitioning from a constrained state in the nanopore to a quenched state on the polynucleotide in free solution. As mentioned above, during this transition interval or period the fluorescent label is capable of emitting a detectable fluorescent signal indicative of the nucleotide it is attached to.

[0052] In some embodiments, the invention includes an application of the discovery that fluorescent labels and nanopores may be selected so that during translocation of a polynucleotide through a nanopore fluorescent labels attached to monomers are forced into a constrained state in which they are incapable (or substantially incapable) of producing a detectable fluorescent signal. In some embodiments, nanopores are selected that have a bore, or lumen, with a diameter in the range of from 1 to 4 nm; in other embodiments, nanopores are selected that have a bore or lumen with a diameter in the range of from 2 to 3 nm. In some embodiments, such bore diameters are provided by a protein nanopore. In some embodiments, such nanopores are used to force fluorescent labels into a constrained state in accordance with the invention, so that whenever a fluorescent label exits a nanopore, it transitions from being substantially incapable of generating a fluorescent signal to being detectable and identifiable by a fluorescent signal it can be induced to emit. Thus, fluorescent labels attached to each of a sequence of monomers of a polynucleotide may be detected in sequence as they suddenly generate a fluorescent signal in a region immediately adjacent to a nanopore exit (a "transition zone" or "transition volume" or "detection zone"). In some embodiments, organic fluorescent dyes are used as fluorescent labels with nanopores of the above diameters. In some embodiments, at least one such organic fluorescent dye is selected from the set consisting of xanthene dyes, rhodamine dyes and cyanine dyes. Some embodiments for determining a monomer sequence of a polynucleotide may be carried out with the following steps: (a) translocating a polynucleotide through a nanopore, wherein monomers of the polynucleotide are labeled with fluorescent labels wherein the nanopore constrains fluorescent labels within its bore into a constrained state such that substantially no detectable fluorescent signal is generated therein; (b) exciting the fluorescent label of each monomer upon exiting the nanopore; (c) measuring a fluorescent signal in a detection zone generated by the exiting fluorescent label to identify the monomer to which the fluorescent label is attached; (d) quenching fluorescent signals from excited fluorescent labels outside of the detection zone, and (d)

determining a monomer sequence of the polynucleotide from a sequence of fluorescent signals. In further embodiments, fluorescent labels are acceptors of a FRET pair and one or more donors of the FRET pair are attached to the nanopore within a FRET distance of the exit.

[0053] In some embodiments, "substantially quenched" as used above means a fluorescent label generates a fluorescent signal at least thirty percent reduced from a signal generated under the same conditions, but without adjacent mutually quenching labels. In some embodiments, "substantially quenched" as used above means a fluorescent label generates a fluorescent signal at least fifty percent reduced from a signal generated under the same conditions, but without adjacent mutually quenching labels.

[0054] In some embodiments, a nucleotide sequence of a target polynucleotide is determined by carrying out four separate reactions in which copies of the target polynucleotide have each of its four different kinds of nucleotide (A, C, G and T) labeled with a single fluorescent label. In a variant of such embodiments, a nucleotide sequence of a target polynucleotide is determined by carrying out four separate reactions in which copies of the target polynucleotide have each of its four different kinds of nucleotide (A, C, G and T) labeled with one fluorescent label while at the same time the other nucleotides on the same target polynucleotide are labeled with a second fluorescent label. For example, if a first fluorescent label is attached to A's of the target polynucleotide in a first reaction, then a second fluorescent label is attached to C's, G's and T's (i.e. to the "not-A" nucleotides) of the target polynucleotides in the first reaction. Likewise, in continuance of the example, in a second reaction, the first label is attached to C's of the target polynucleotide and the second fluorescent label is attached to A's, G's and T's (i.e. to the "not-C" nucleotides) of the target polynucleotide. And so on, for nucleotides G and T.

[0055] The same labeling scheme may be expressed in terms of conventional terminology for subsets of nucleotide types; thus, in the above example, in a first reaction, a first fluorescent label is attached to A's and a second fluorescent label is attached to B's; in a second reaction, a first fluorescent label is attached to C's and a second fluorescent label is attached to D's; in a third reaction, a first fluorescent label is attached to G's and a second fluorescent label is attached to H's; and in a fourth reaction, a first fluorescent label is attached to T's and a second fluorescent label is attached to V's.

[0056] In some embodiments, a polymer, such as a polynucleotide or peptide, may be labeled with a single fluorescent label attached to a single kind of monomer, for example, every T (or substantially every T) of a polynucleotide is labeled with a fluorescent label, e.g. a cyanine dye. In such embodiments, a collection, or sequence, of fluorescent signals from the polynucleotide may form a signature or fingerprint for the particular polynucleotide. In some such embodiments, such fingerprints may or may not provide enough information for a sequence of monomers to be determined.

[0057] In some embodiments, a feature of the invention is the labeling of substantially all monomers of a polynucleotide analyte with fluorescent dyes or labels that are members of a mutually quenching set. The use of the term "substantially all" in reference to labeling polynucleotide analytes is to acknowledge that chemical and enzymatic labeling techniques are typically less than 100 percent efficient. In some embodiments, "substantially all" means at least 80 percent of all monomer have fluorescent labels attached. In other embodiments, "substantially all" means at least 90 percent of all monomer have fluorescent labels attached. In other embodiments, "substantially all" means at least 95 percent of all monomer have fluorescent labels attached. Mutually quenching sets of fluorescent dyes have the following properties: (i) each member quenches fluorescence of every member (for example, by FRET or by static or contact mechanisms), and (ii) each member generates a distinct fluorescent signal when excited and when in a non-quenched state. That is, if a mutually quenching set consists of two dyes, D1 and D2, then (i) D1 is self-quenched (e.g. by contact quenching with another D1 molecule) and it is quenched by D2 (e.g. by contact quenching) and (ii) D2 is self-quenched (e.g. by contact quenching with another D2 molecule) and it is quenched by D1 (e.g. by contact quenching). Guidance for selecting fluorescent dyes or labels for mutually quenching sets may be found in the following references, which are incorporated herein by reference: Johansson, Methods in Molecular Biology, 335: 17-29 (2006); Marras et al, Nucleic Acids Research, 30: e122 (2002); and the like. In some embodiments, members of a mutually quenching set comprise organic fluorescent dyes that components or moieties capable of stacking interactions, such as aromatic ring structures. Exemplary mutually quenching sets of fluorescent dyes, or labels, may be selected from rhodamine dyes, fluorescein dyes and cyanine dyes. In one embodiment, a mutually quenching set may comprise the rhodamine dye, TAMRA, and the fluorescein dye, FAM. In another embodiment, mutually quenching sets of fluorescent dyes may be formed by selecting two or more dyes from the group consisting of Oregon Green 488, Fluorescein-EX, fluorescein isothiocyanate, Rhodamine Red-X, Lissamine rhodamine B, Calcein, Fluorescein, Rhodamine, one or more BODIPY dyes, Texas Red, Oregon Green 514, and one or more Alexa Fluors. Respresentative BODIPY dyes include BODIPY FL, BODIPY R6G, BODIPY TMR, BODIPY 581/591, BODIPY TR, BODIPY 630/650 and BODIPY 650/665. Representative Alexa Fluors include Alexa Fluor 350, 405, 430, 488, 500, 514, 532, 546, 555, 568, 594, 610, 633, 635, 647, 660, 680, 700, 750 and 790.

[0058] As above, in some embodiments, a monomer sequence of a target polynucleotide is determined by carrying out separate reactions (one for each kind of monomer) in which copies of the target polynucleotide have each different kind of monomer labeled with a mutually- or self-quenching fluorescent label. In other embodiments, a monomer sequence of a target polynucleotide is determined by carrying out separate reactions (one for each kind of monomer) in which copies of the target

polynucleotide have each different kind of monomer labeled with a different mutually quenching fluorescent label selected from the same mutually quenching set. In embodiments in which a mutually quenching set contains only two dyes, then a selected monomer (say, monomer X) is labeled with a first mutually quenching dye and every other kind of monomer (i.e., not-monomer X) is labeled with a second mutually quenching dye from the same set. Thus, steps of the embodiment generate a sequence of two different fluorescent signals, one indicating monomer X and another indicating not-monomer X.

[0059] In some embodiments, a single fluorescent label (for example, attached to a single kind of monomer in a polynucleotide comprising multiple kinds of monomers) may be used that is self-quenching when attached to adjacent monomers (of the same kind) on a polynucleotide, such as adjacent nucleotides of a polynucleotide. Exemplary self-quenching fluorescent labels include, but are not limited to, Oregon Green 488, fluorescein-EX, FITC, Rhodamine Red-X, Lissamine rhodamine B, calcein, fluorescein, rhodamine, BODIPYS, and Texas Red, e.g. which are disclosed in Molecular Probes Handbook, 11th Edition (2010).

Embodiments Employing Quenching Agents

[0060] Figs. 6A-6C illustrate different embodiments of the invention corresponding to where quenching agents are applied in a nanopore device: trans chamber only (Fig. 6A), cis chamber only (Fig. 6B), or both cis and trans chambers (Fig. 6C). In Fig. 6A, labeled polynucleotide (600) is illustrated translocating nanopore (606) of solid phase membrane (608) from cis chamber (602) to trans chamber (604). Immersed in trans chamber (604) are non-fluorescent quenching agents (605) designated by "Q". Quenching agents of the invention are soluble under translocation conditions for labeled polynucleotide (600), and under the same conditions, quenching agents bind to single stranded polynucleotides, such as (600), without substantial sequence specificity. As explained more fully below, a large variety of non-fluorescent quenching agents are available for use with the invention, which include derivatives of many well-known organic dyes, such as asymmetric cyanine dyes, as well as conjugates of such compounds and oligonucleotides and/or analogs thereof. In this embodiment, selection of the type and concentration of quenching agent and the translocation speed define detection zone (610). In some embodiments, "detection zone" means a region or volume (which may be contiguous or non-contiguous) from which fluorescent signals are collected to form the raw data from which information, such as sequence information, about a labeled polynucleotide is determined. Fluorescent labels in trans chamber (604) outside of detection zone (610) are substantially quenched by quenching agents (605) bound to the portion of labeled polynucleotide (600) in trans chamber (604). In some embodiments, quenching agents comprise an oligonucleotide or analog conjugated to one or more quenching moieties based on organic dyes as described more fully below. Embodiments of Fig. 6A

may be employed when, for example, solid phase membrane (608) is or comprises an opaque layer so that fluorescent labels in cis chamber (602) are substantially non-excited.

[0061] Fig. 6B shows substantially the same elements as those in Fig. 6A with the exception that quenching agents (605) are disposed in cis chamber (602). This configuration may be desirable under circumstances where undesired evanescent waves, or like non-radiative light energy, extend to cis chamber (602) and excite fluorescent labels which generate fluorescent signals that are collected. Quenching agents (605) that bind to labeled polynucleotide (600) in cis chamber (602) reduce or eliminate such fluorescent signals. In some embodiments, quenching agents (605) and cross-section of nanopore (606) are selected so that quenching agents (605) are excluded from translocating through nanopore (606). In some embodiments, this may be achieved by using protein nanopore α-hemolysin and quenching agents comprising conjugates of oligonucleotides or analogs thereof and one or more quenching compounds, as described more fully below.

[0062] Fig. 6C illustrates an embodiment where quenching agents (605) are present in both cis chamber (602) and trans chamber (604), which provides the advantages described for the embodiments of both Figs. 6A and 6B.

[0063] Fig. 7 illustrates an embodiment which includes the following elements: protein nanopore (700) disposed in lipid bilayer (702); epi-illumination of fluorescent labels with opaque layer (708) in solid phase membrane (706) to prevent or reduce background fluorescence; and quenching agents (710) disposed in trans chamber (726). As above, polynucleotide (720) with fluorescently labeled nucleotides (labels being indicated by "f", as with (722)) is translocated through nanopore (700) from cis chamber (724) to trans chamber (726). Oligonucleotide quenchers (710) are disposed in trans chamber (726) under conditions (e.g. concentration, temperature, salt concentration, and the like) that permits hybridization of oligonucleotide quenchers (728) to portions of polynucleotide (720) emerging from nanopore (700). Nanopore (700) may be selected so that signals from fluorescent labels are suppressed during transit of the nanopore as described in Huber et al, U.S. patent publication US 2016/0076091, which is incorporated herein by reference. Thus, when labeled nucleotides emerge from nanopore (700) in region (728) they become unsuppressed and capable of generating a signal. With most if not all forms of direct illumination (e.g. non-FRET) such emerged labels would continue to emit fluorescence as they travel further into trans chamber (726), thereby contributing greatly to a collected signal. With quenching agents in trans chamber (726) that bind to the emerging polynucleotide, such emissions can be significantly reduced and can define detection zone (728) from which collected signals can be analyzed to give nucleotide sequence information about polynucleotide (720). In some embodiments, a fluorescent signal from a single fluorescent label is detected from detection zone (728) during a detection period as the labeled polynucleotide moves through the

detection zone. In other embodiments, a plurality of fluorescent signals is collected from a plurality of fluorescent labels in detection zone (728) during a predetermined time period. In some embodiments, such detection period is less than 1 msec, or less than 0.1 msec, or less than 0.01 msec. In some embodiments, such detection perior is at least 0.01 msec, or at least 0.1 msec, or at least 0.5 msec.

[0064] Quenching agents of the invention comprise any compound (or set of compounds) that under nanopore sequencing conditions is (i) substantially non-fluorescent, (ii) binds to single stranded nucleic acids, particularly single stranded DNA, and (iii) absorbs excitation energy from other molecules non-radiatively and releases it non-radiatively. In some embodiments, quenching agents further bind non-covalently to single stranded DNA. A large variety of quenching compounds are available for use with the invention including, but not limited to, non-fluorescent derivatives of common synthetic dyes such as cyanine and xanthene dyes, as described more fully below. Guidance in selecting quenching compounds may be found in U.S. patents 6,323,337; 6,750,024 and like references, which are incorporated herein by reference.

[0065] In some embodiments, a quenching agent may be a single stranded DNA binding dye that has been covalently modified with a heavy atom that is known to quench fluorescence (such as bromine or iodine), or covalently modified with other groups known to quench fluorescence, such as a nitro group or a azo group. An example of dye that is known to bind single stranded DNA is Sybr Green (Zipper et al, (2004), Nucleic Acids Research. 32 (12)). Incorporation of a nitro, bromine, iodine, and/or azo groups into the cynanine Sybr Green structure provides a single stranded DNA binding group moiety that will quench fluorescent labels that might be present on a DNA.

[0066] In some embodiments, quenching agents comprise a binding moiety and one or more quenching moieties. Binding moieties may include any compound that binds to single stranded nucleic acids without substantial sequence specificity. Binding moieties may comprise peptides or oligonucleotides or analogs of either having modified linkages and/or monomers. Oligonucleotides and their analogs may provide binding to polynucleotides via duplex formation or via non-base paired aptameric binding. In some embodiments, binding moieties comprise an oligonucleotide or analog thereof having a length in the range of from 6 to 60 nucleotides. Such oligonucleotides or analogs may be conjugated to one quenching moiety or to a plurality of quenching moieties. In some embodiments, the plurality of quenching moieties conjugated to each oligonucleotide or analog is 2 or 3. Quenching moieties conjugated to a binding moiety may be the same or different. In some embodiments, whenever a binding moiety is an oligonucleotide or analog, two quenching moieties are conjugated thereto, one at a 5' end and one at a 3' end of the oligonucleotide. Oligonucleotides or analogs having

from 2 to 3 quenching moieties may be synthesized using conventional linkage and synthetic chemistries, for example, as disclosed in the references cited herein.

[0067] Oligonucleotides or analogs may be provided as a single species or they may be provided as mixtures of a plurality of oligonucleotides or analogs with different sequences, and therefore, different binding specificities. In some embodiments, oligonucleotides or analogs are random sequence polymers; that is, they are provided as mixtures of every possible sequence of a given length. For example, such oligonucleotides or analogs may be represented by the formulas, "NNNNNN" for 6-mers, or "NNNNNNNN" for 8-mers, wherein N may be A, C, G or T, or an analog thereof.

[0068] "Analogs" in reference to oligonucleotides means an oligonucleotide that contains one or more nucleotide analogs. As described in the definition section, a "nucleotide analog" is a nucleotide that may have a modified linkage moiety, sugar moiety or base moiety. Exemplary oligonucleotide analogs that may be used with the invention include, but are not limited to, peptide nucleic acids (PNAs), locked nucleic acids (LNAs)(2'-O-methyl RNA), phosphorothioate oligonucleotides, bridged nucleic acids (BNAs), or the like.

[0069] In some embodiments, oligonucleotide binding moieties comprise universal bases; that is, they contain one or more nucleotide analogs that can replace any of the four natural nucleotides without destabilizing base-pair interactions. Nucleotide analogs having universal base properties are described in Loakes, Nucleic Acids Research, 29(12): 2437-2447 (2001), which is incorporated herein by reference. In some embodiments, oligonucleotide binding moieties comprise 2'-deoxyinosine, 7-deaza-2'-deoxyinosine, 2-aza-2'-deoxyinosine, 3-nitropyrrole nucleotides, 5-nitroindole nucleotides, or the like.

[0070] In some embodiments, quenching agents may comprise a combination of two or more compounds that act together to quench undesired fluorescent signals of a single stranded labeled polynucleotide. For example, a quenching agent may comprise an oligonucleotide (e.g., polydeoxyinosine) that may form a duplex with the labeled polynucleotide and separately a double stranded intercalator that is a quencher. Thus, whenever the polydeoxyinosine binds to a labeled polynucleotide, the quenching intercalator binds to the resulting duplex and quenches fluorescent signals from the polynucleotide.

[0071] Any synthetic dye that can detectably quench fluorescent signals of the fluorescent labels of a labeled polynucleotide is an acceptable quenching moiety for the purposes of the invention. Specifically, as used in the invention, the quenching moieties possess an absorption band that exhibits at least some spectral overlap with an emission band of the fluorescent labels on a labeled polynucleotide. This overlap may occur with emission of the fluorescent label (donor) occurring at a lower or even higher wavelength emission maximum than the maximal absorbance wavelength of the

quenching moiety (acceptor), provided that sufficient spectral overlap exists. Energy transfer may also occur through transfer of emission of the donor to higher electronic states of the acceptor. One of ordinary skill in the art determines the utility of a given quenching moiety by examination of that dye's excitation bands with respect to the emission spectrum of the fluorescent labels being used.

[0072] Typically, fluorescence quenching in the invention occurs through Fluorescence Resonance Energy Transfer (FRET or through the formation of charge transfer complexes) between a fluorescent label and a quenching moiety of the invention. The spectral and electronic properties of the donor and acceptor compounds have a strong effect on the degree of energy transfer observed, as does the separation distance between the fluorescent labels on the labeled polynucleotide and the quenching moiety. As the separation distance increases, the degree of fluorescence quenching decreases.

[0073] A quenching moiety may be optionally fluorescent, provided that the maximal emission wavelength of the dye is well separated from the maximal emission wavelength of the fluorescent labels when bound to labeled polynucleotides. Preferably, however, the quenching moiety is only dimly fluorescent, or is substantially non-fluorescent, when covalently conjugated to a oligonucleotide or analog. Substantially non-fluorescent, as used herein, indicates that the fluorescence efficiency of the quenching moiety in an assay solution as described for any of the methods herein is less than or equal to 5 percent, preferably less than or equal to 1 percent. In other embodiments, the covalently bound quenching moiety exhibits a quantum yield of less than about 0.1, more preferably less than about 0.01. In some embodiments, the fluorescence of fluorescent labels associated with a quenching oligonucleotide of the invention is quenched more than 50% relative to the same oligonucleotide associated with the same fluorescent labels in the absence of the covalently bound quenching moiety. In another embodiment, the fluorescent labels are quenched more than 90% relative to the unlabeled oligonucleotide. In yet another embodiment, the nucleic acid stains are quenched more than 95% relative to the unlabeled oligonucleotide.

[0074] In some embodiments, a quenching moiety may be a pyrene, an anthracene, a naphthalene, an acridine, a stilbene, an indole or benzindole, an oxazole or benzoxazole, a thiazole or benzothiazole, a 4-amino-7-nitrobenz-2-oxa-1,3-diazole (NBD), a cyanine, a carbocyanine, a carbostyryl, a porphyrin, a salicylate, an anthranilate, an azulene, a perylene, a pyridine, a quinoline, a coumarin (including hydroxycoumarins and aminocoumarins and fluorinated and sulfonated derivatives thereof (as described in U.S. Pat. No. 5,830,912 to Gee et al. (1998) and U.S. Pat. No. 5,696,157 to Wang et al. (1997), incorporated by reference), a polyazaindacene (e.g. U.S. Pat. No. 4,774,339 to Haugland, et al. (1988); U.S. Pat. No. 5,187,288 to Kang, et al. (1993); U.S. Pat. No. 5,248,782 to Haugland, et al. (1993); U.S. Pat. No. 5,274,113 to Kang, et al. (1993); 5,433,896 to Kang, et al.(1995); U.S. Pat. No. 6,005,113 to Wu et al. (1999), all incorporated by reference), a xanthene, an oxazine or a benzoxazine,

a carbazine (U.S. Pat. No. 4,810,636 to Corey (1989), incorporated by reference), or a phenalenone or benzphenalenone (U.S. Pat. No. 4,812,409 Babb et al. (1989), incorporated by reference).

[0075] In other embodiments, quenching moieties that are substantially non-fluorescent dyes include in particular azo dyes (such as DABCYL or DABSYL dyes and their structural analogs), triarylmethane dyes such as malachite green or phenol red, 4',5z-diether substituted fluoresceins (U.S. Pat. No. 4,318,846 (1982)), or asymmetric cyanine dye quenchers (PCT Int. App. WO 99 37,717 (1999)).

[0076] In embodiments where the quenching moiety is a xanthene, the synthetic dye is optionally a fluorescein, a rhodol (U.S. Pat. No. 5,227,487 to Haugland, et al. (1993), incorporated by reference), or a rhodamine. As used herein, fluorescein includes benzo- or dibenzofluoresceins, seminaphthofluoresceins, or naphthofluoresceins. Similarly, as used herein rhodol includes seminaphthorhodafluors (U.S. Pat. No. 4,945,171 to Haugland, et al. (1990), incorporated by reference). Xanthenes include fluorinated derivatives of xanthene dyes (Int. Publ. No. WO 97/39064, Molecular Probes, Inc. (1997), incorporated by reference), and sulfonated derivatives of xanthene dyes (Int. Publ. No. WO 99/15517, Molecular Probes, Inc. (1999), incorporated by reference). As used herein, oxazines include resorufms, aminooxazinones, diaminooxazines, and their benzo-substituted analogs.

[0077] In further embodiments, the quenching moiety is an substantially nonfluorescent derivative of 3- and/or 6-amino xanthene that is substituted at one or more amino nitrogen atoms by an aromatic or heteroaromatic ring system, e.g. as described in U.S. patent 6,399,392, which is incorporated herein by reference. These quenching dyes typically have absorption maxima above 530 nm, have little or no observable fluorescence and efficiently quench a broad spectrum of luminescent emission, such as is emitted by chemilumiphores, phosphors, or fluorophores. In one embodiment, the quenching dye is a substituted rhodamine. In another embodiment, the quenching compound is a substituted rhodol.

[0078] In still other embodiments, a quenching moiety may comprise one or more non-fluorescent quenchers known as Black Hole Quenchers™ compounds (BHQs) described in the following patents, which are incorporated herein by reference: 7,019,129; 7,109,312; 7,582,432; 8,410,025; 8,440,399; 8,633,307; 8,946,404; 9,018,369; or 9,139,610.

Additional quenching moieties are disclosed in the following, which are incorporated herein by reference: U.S. patents 6,699,975; 6,790,945; and 8,114,979.

Labels for Nanopores and Polymers

[0079] In some embodiments, a nanopore may be labeled with one or more quantum dots. In particular, in some embodiments, one or more quantum dots may be attached to a nanopore, or attached

to a solid phase support adjacent to (and within a FRET distance of an entrance or exit of a nanopore), and employed as donors in FRET reactions with acceptors on analytes. Such uses of quantum dots are well known and are described widely in the scientific and patent literature, such as, in U.S. patents 6,252,303; 6,855,551; 7,235,361; and the like, which are incorporated herein by reference.

[0080] One example of a Quantum dot which may be utilized as a pore label is a CdTe quantum dot which can be synthesized in an aqueous solution. A CdTe quantum dot may be functionalized with a nucleophilic group such as primary amines, thiols or functional groups such as carboxylic acids. A CdTe quantum dot may include a mercaptopropionic acid capping ligand, which has a carboxylic acid functional group that may be utilized to covalently link a quantum dot to a primary amine on the exterior of a protein pore. The cross-linking reaction may be accomplished using standard cross-linking reagents (homo-bifunctional as well as hetero-bifunctional) which are known to those having ordinary skill in the art of bioconjugation. Care may be taken to ensure that the modifications do not impair or substantially impair the translocation of a nucleic acid through the nanopore. This may be achieved by varying the length of the employed crosslinker molecule used to attach the donor label to the nanopore.

[0081] For example, the primary amine of the lysine residue 131 of the natural alpha hemolysin protein (Song, L. et al., Science 274, (1996): 1859-1866) may be used to covalently bind carboxy modified CdTe Quantum dots via 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride/ N-hydroxysulfosuccinimide (EDC/NHS) coupling chemistry. Alternatively, amino acid 129 (threonine) may be exchanged into cysteine. Since there is no other cysteine residue in the natural alpha hemolysin protein the thiol side group of the newly inserted cysteine may be used to covalently attach other chemical moieties.

[0082] A biological polymer, e.g., a nucleic acid molecule or polymer, may be labeled with one or more acceptor labels. For a nucleic acid molecule, each of the four nucleotides or building blocks of a nucleic acid molecule may be labeled with an acceptor label thereby creating a labeled (e.g., fluorescent) counterpart to each naturally occurring nucleotide. The acceptor label may be in the form of an energy accepting molecule which can be attached to one or more nucleotides on a portion or on the entire strand of a converted nucleic acid.

[0083] A variety of methods may be utilized to label the monomers or nucleotides of a nucleic acid molecule or polymer. A labeled nucleotide may be incorporated into a nucleic acid during synthesis of a new nucleic acid using the original sample as a template ("labeling by synthesis"). For example, the labeling of nucleic acid may be achieved via PCR, whole genome amplification, rolling circle amplification, primer extension or the like or via various combinations and extensions of the above methods known to persons having ordinary skill in the art.

[0084] A label may comprise a reactive group such as a nucleophile (amines, thiols etc.). Such nucleophiles, which are not present in natural nucleic acids, can then be used to attach fluorescent labels via amine or thiol reactive chemistry such as NHS esters, maleimides, epoxy rings, isocyanates etc. Such nucleophile reactive fluorescent dyes (i.e. NHS-dyes) are readily commercially available from different sources. An advantage of labeling a nucleic acid with small nucleophiles lies in the high efficiency of incorporation of such labeled nucleotides when a "labeling by synthesis" approach is used. Bulky fluorescently labeled nucleic acid building blocks may be poorly incorporated by polymerases due to steric hindrance of the labels during the polymerization process into newly synthesized DNA.

[0085] Whenever two or more mutually quenching dyes are used, such dyes may be attached to DNA using orthogonal attachment chemistries. For example, NHS esters can be used to react very specifically with primary amines or maleimides will react with thiol groups. Either primary amines (NH2) or thiol (SH) modified nucleotides are commercially available. These relatively small modifications are readily incorporated in a polymerase mediated DNA synthesis and can be used for subsequent labeling reactions using either NHS or maleimide modified dyes. Guidance for selecting and using such orthogonal linker chemistries may be found in Hermanson (cited above).

[0086] Additional orthogonal attachment chemistries for typical attachment positions include Huisgen-type cycloaddition for a copper-catalyzed reaction and an uncatalyzed reaction; alkene plus nitrile oxide cycloaddition, e.g. as disclosed in Gutsmiedl et al, Org. Lett., 11: 2405-2408 (2009); Diels-Alder cycloaddition, e.g. disclosed in Seelig et al, Tetrahedron Lett., 38: 7729-7732 (1997); carbonyl ligation, e.g. as disclosed in Casi et al, J. Am. Chem. Soc., 134: 5887-5892 (2012); Shao et al J. Am. Chem. Soc., 117: 3893-3899 (1995); Rideout, Science, 233: 561-563 (1986); Michael addition, e.g. disclosed in Brinkley, Bioconjugate Chemistry, 3: 2-13 (1992); native chemical ligation, e.g. disclosed in Schuler et al, Bioconjugate Chemistry, 13: 1039-1043 (2002); Dawson et al, Science, 266: 776-779 (1994); or amide formation via an active ester, e.g. disclosed in Hermanson (cited above).

Definitions

[0087] "Evanescent field" means a non-propagating electromagnetic field; that is, it is an electromagnetic field in which the average value of the Poynting vector is zero.

[0088] "FRET" or "Förster, or fluorescence, resonant energy transfer" means a non-radiative dipole-dipole energy transfer mechanism from an excited donor fluorophore to an acceptor fluorophore in a ground state. The rate of energy transfer in a FRET interaction depends on the extent of spectral overlap of the emission spectrum of the donor with the absorption spectrum of the acceptor, the quantum yield of the donor, the relative orientation of the donor and acceptor transition dipoles, and the distance between the donor and acceptor molecules, Lakowitz, Principles of Fluorescence

Spectroscopy, Third Edition (Springer, 2006). FRET interactions of particular interest are those which result a portion of the energy being transferred to an acceptor, in turn, being emitted by the acceptor as a photon, with a frequency lower than that of the light exciting its donor (i.e. a "FRET signal"). "FRET distance" means a distance between a FRET donor and a FRET acceptor over which a FRET interaction can take place and a detectable FRET signal produced by the FRET acceptor.

[0089] "Kit" refers to any delivery system for delivering materials or reagents for carrying out a method of the invention. In the context of reaction assays, such delivery systems include systems that allow for the storage, transport, or delivery of reaction reagents (e.g., fluorescent labels, such as mutually quenching fluorescent labels, fluorescent label linking agents, enzymes, etc. in the appropriate containers) and/or supporting materials (e.g., buffers, written instructions for performing the assay etc.) from one location to another. For example, kits include one or more enclosures (e.g., boxes) containing the relevant reaction reagents and/or supporting materials. Such contents may be delivered to the intended recipient together or separately. For example, a first container may contain an enzyme for use in an assay, while a second or more containers contain mutually quenching fluorescent labels.

[0090] "Nanopore" means any opening positioned in a substrate that allows the passage of analytes through the substrate in a predetermined or discernable order, or in the case of polymer analytes, passage of their monomeric units through the substrate in a pretermined or discernible order. In the latter case, a predetermined or discernible order may be the primary sequence of monomeric units in the polymer. Examples of nanopores include proteinaceous or protein based nanopores, synthetic or solid state nanopores, and hybrid nanopores comprising a solid state nanopore having a protein nanopore embedded therein. A nanopore may have an inner diameter of 1-10 nm or 1-5 nm or 1-3 nm. Examples of protein nanopores include but are not limited to, alpha-hemolysin, voltage-dependent mitochondrial porin (VDAC), OmpF, OmpC, MspA and LamB (maltoporin), e.g. disclosed in Rhee, M. et al., Trends in Biotechnology, 25(4) (2007): 174-181; Bayley et al (cited above); Gundlach et al, U.S. patent publication 2012/0055792; and the like, which are incorporated herein by reference. Any protein pore that allows the translocation of single nucleic acid molecules may be employed. A nanopore protein may be labeled at a specific site on the exterior of the pore, or at a specific site on the exterior of one or more monomer units making up the pore forming protein. Pore proteins are chosen from a group of proteins such as, but not limited to, alpha-hemolysin, MspA, voltage-dependent mitochondrial porin (VDAC), Anthrax porin, OmpF, OmpC and LamB (maltoporin). Integration of the pore protein into the solid state hole is accomplished by attaching a charged polymer to the pore protein. After applying an electric field the charged complex is electrophoretically pulled into the solid state hole. A synthetic nanopore, or solid-state nanopore, may be created in various forms of solid substrates, examples of which include but are not limited to silicones (e.g. Si_3N_4, SiO_2), metals, metal

oxides (e.g. Al2O3) plastics, glass, semiconductor material, and combinations thereof. A synthetic nanopore may be more stable than a biological protein pore positioned in a lipid bilayer membrane. A synthetic nanopore may also be created by using a carbon nanotube embedded in a suitable substrate such as but not limited to polymerized epoxy. Carbon nanotubes can have uniform and well-defined chemical and structural properties. Various sized carbon nanotubes can be obtained, ranging from one to hundreds of nanometers. The surface charge of a carbon nanotube is known to be about zero, and as a result, electrophoretic transport of a nucleic acid through the nanopore becomes simple and predictable (Ito, T. et al., Chem. Commun. 12 (2003): 1482-83). The substrate surface of a synthetic nanopore may be chemically modified to allow for covalent attachment of the protein pore or to render the surface properties suitable for optical nanopore sequencing. Such surface modifications can be covalent or non-covalent. Most covalent modification include an organosilane deposition for which the most common protocols are described:1) Deposition from aqueous alcohol. This is the most facile method for preparing silylated surfaces. A 95% ethanol-5% water solution is adjusted to pH 4.5-5.5 with acetic acid. Silane is added with stirring to yield a 2% final concentration. After hydrolysis and silanol group formation the substrate is added for 2-5min. After rinsed free of excess materials by dipping briefly in ethanol. Cure of the silane layer is for 5-10min at 110 degrees Celsius. 2) Vapor Phase Deposition. Silanes can be applied to substrates under dry aprotic conditions by chemical vapor deposition methods. These methods favor monolayer deposition. In closed chamber designs, substrates are heated to sufficient temperature to achieve 5mm vapor pressure. Alternatively, vacuum can be applied until silane evaporation is observed. 3) Spin-on deposition. Spin-on applications can be made under hydrolytic conditions which favor maximum functionalization and polylayer deposition or dry conditions which favor monolayer deposition. In some embodiments, single nanopores are employed with methods of the invention. In other embodiments, a plurality of nanopores are employed. In some of the latter embodiments, a plurality of nanopores is employed as an array of nanopores, usually disposed in a planar substrate, such as a solid phase membrane. Nanopores of a nanopore array may be spaced regularly, for example, in a rectilinear pattern, or may be spaced randomly. In a preferred embodiment, nanopores are spaced regularly in a rectilinear pattern in a planar solid phase substrate.

[0091] "Nanostructure" (used interchangeably with "nanoscale structure" and "nanoscale feature") means a structure that has at least one dimension within a range of a few nanometers to several hundred nanometers, for example, from 1 to 1000 nanometers. In some applications, such range is from 2 to 500 nanometers; in other applications, such range is from 3 to 500 nanometers. The shape and geometry of nanostructures may vary widely and include, but are not limited to, nanopores, nanowells, nanoparticles, and any other convenient shapes particularly suitable for carrying out sequences of reactions. In some embodiments, nanostructures may be protein nanopores operationally associated

with a solid phase membrane. Some nanostructures, such as, nanopores and nanowells, may be formed in a larger common substrate, such as a solid phase membrane, or other solid, to form arrays of nanopores or nanowells. Nanostructures of particular interest are those capable of supporting or containing a chemical, physical (e.g. FRET), enzymatic and/or binding reaction or a sequence of such reactions. In some embodiments, a nanostructure, such as a nanowell, encloses a volume that is less than one nanoliter (10×−9 liter), less than one picoliter, or less than one femtoliter. In other embodiments, each of the individual nanowells provides a volume that is less than 1000 zeptoliters, 100 zeptoliters, 80 zeptoliters, or less than 50 zeptoliters, or less than 1 zeptoliter, or even less than 100 yactoliters. In some embodiments, nanowells comprise zero mode waveguides.

[0092] "Peptide," "peptide fragment," "polypeptide," "oligopeptide," or "fragment" in reference to a peptide are used synonymously herein and refer to a compound made up of a single unbranched chain of amino acid residues linked by peptide bonds. Amino acids in a peptide or polypeptide may be derivatized with various moieties, including but not limited to, polyethylene glycol, dyes, biotin, haptens, or like moieties. The number of amino acid residues in a protein or polypeptide or peptide may vary widely; however, in some embodiments, protein or polypeptides or peptides referred to herein may have 2 from to 70 amino acid residues; and in other embodiments, they may have from 2 to 50 amino acid residues. In other embodiments, proteins or polypeptides or peptides referred to herein may have from a few tens of amino acid residues, e.g. 20, to up to a thousand or more amino acid residues, e.g. 1200. In still other embodiments, proteins, polypeptides, peptides, or fragments thereof, may have from 10 to 1000 amino acid residues; or they may have from 20 to 500 amino acid residues; or they may have from 20 to 200 amino acid residues.

[0093] "Polymer" means a plurality of monomers connected into a linear chain. Usually, polymers comprise more than one type of monomer, for example, as a polynucleotide comprising A's, C's, G's and T's, or a polypeptide comprising more than one kind of amino acid. Monomers may include without limitation nucleosides and derivatives or analogs thereof and amino acids and derivatives and analogs thereof. In some embodiments, polymers are polynucleotides, whereby nucleoside monomers are connected by phosphodiester linkages, or analogs thereof.

[0094] "Polynucleotide" or "oligonucleotide" are used interchangeably and each mean a linear polymer of nucleotide monomers. Monomers making up polynucleotides and oligonucleotides are capable of specifically binding to a natural polynucleotide by way of a regular pattern of monomer-to-monomer interactions, such as Watson-Crick type of base pairing, base stacking, Hoogsteen or reverse Hoogsteen types of base pairing, or the like. Such monomers and their internucleosidic linkages may be naturally occurring or may be analogs thereof, e.g. naturally occurring or non-naturally occurring analogs. Non-naturally occurring analogs may include PNAs, phosphorothioate internucleosidic

linkages, bases containing linking groups permitting the attachment of labels, such as fluorophores, or haptens, and the like. Whenever the use of an oligonucleotide or polynucleotide requires enzymatic processing, such as extension by a polymerase, ligation by a ligase, or the like, one of ordinary skill would understand that oligonucleotides or polynucleotides in those instances would not contain certain analogs of internucleosidic linkages, sugar moieties, or bases at any or some positions. Polynucleotides typically range in size from a few monomeric units, e.g. 5-40, when they are usually referred to as "oligonucleotides," to several thousand monomeric units. Whenever a polynucleotide or oligonucleotide is represented by a sequence of letters (upper or lower case), such as "ATGCCTG," it will be understood that the nucleotides are in 5'→3' order from left to right and that "A" denotes deoxyadenosine, "C" denotes deoxycytidine, "G" denotes deoxyguanosine, and "T" denotes thymidine, "I" denotes deoxyinosine, "U" denotes uridine, unless otherwise indicated or obvious from context. Unless otherwise noted the terminology and atom numbering conventions will follow those disclosed in Strachan and Read, Human Molecular Genetics 2 (Wiley-Liss, New York, 1999). Usually polynucleotides comprise the four natural nucleosides (e.g. deoxyadenosine, deoxycytidine, deoxyguanosine, deoxythymidine for DNA or their ribose counterparts for RNA) linked by phosphodiester linkages; however, they may also comprise non-natural nucleotide analogs, e.g. including modified bases, sugars, or internucleosidic linkages. It is clear to those skilled in the art that where an enzyme has specific oligonucleotide or polynucleotide substrate requirements for activity, e.g. single stranded DNA, RNA/DNA duplex, or the like, then selection of appropriate composition for the oligonucleotide or polynucleotide substrates is well within the knowledge of one of ordinary skill, especially with guidance from treatises, such as Sambrook et al, Molecular Cloning, Second Edition (Cold Spring Harbor Laboratory, New York, 1989), and like references. Likewise, the oligonucleotide and polynucleotide may refer to either a single stranded form or a double stranded form (i.e. duplexes of an oligonucleotide or polynucleotide and its respective complement). It will be clear to one of ordinary skill which form or whether both forms are intended from the context of the terms usage.

[0095] "Primer" means an oligonucleotide, either natural or synthetic that is capable, upon forming a duplex with a polynucleotide template, of acting as a point of initiation of nucleic acid synthesis and being extended from its 3' end along the template so that an extended duplex is formed. Extension of a primer is usually carried out with a nucleic acid polymerase, such as a DNA or RNA polymerase. The sequence of nucleotides added in the extension process is determined by the sequence of the template polynucleotide. Usually primers are extended by a DNA polymerase. Primers usually have a length in the range of from 14 to 40 nucleotides, or in the range of from 18 to 36 nucleotides. Primers are employed in a variety of nucleic amplification reactions, for example, linear amplification reactions using a single primer, or polymerase chain reactions, employing two or more primers. Guidance for

selecting the lengths and sequences of primers for particular applications is well known to those of ordinary skill in the art, as evidenced by the following references that are incorporated by reference: Dieffenbach, editor, PCR Primer: A Laboratory Manual, 2nd Edition (Cold Spring Harbor Press, New York, 2003).

[0096] "Resolution limited area" is an area of a surface of a nanopore or nanowell array within which individual features or light emission sources cannot be distinguished by an optical signal detection system. Without intending to be limited by theory, such resolution limited area is determined by a resolution limit (also sometimes referred to as a "diffraction limit" or "diffraction barrier") of an optical system. Such limit is determined by the wavelength of the emission source and the optical components and may be defined by $d=\lambda/NA$, where d is the smallest feature that can be resolved, λ is the wavelength of the light and NA is the numerical aperture of the objective lens used to focus the light. Thus, whenever two or more nanopores are within a resolution limited area and two or more optical signals are generated at the respective nanopores, an optical detection system cannot distinguish or determine which optical signals came from which nanopore. In accordance with the invention, a surface of a nanopore array may be partitioned, or subdivided, into non-overlapping regions, or substantially non-overlapping regions, corresponding to resolution limited areas. The size of such subdivisions corresponding to resolution limited areas may depend on a particular optical detection system employed. In some embodiments, whenever light emission sources are within the visible spectrum, a resolution limited area is in the range of from 300 nm^2 to 3.0 μm^2; in other embodiments, a resolution limited area is in the range of from 1200 nm^2 to 0.7 μm^2; in other embodiments, a resolution limited area is in the range of from $3x10^4$ nm^2 to 0.7 μm^2, wherein the foregoing ranges of areas are in reference to a surface of a nanopore or nanowell array. In some embodiments, the visible spectrum means wavelengths in the range of from about 380 nm to about 700 nm.

[0097] "Sequence determination", "sequencing" or "determining a nucleotide sequence" or like terms in reference to polynucleotides includes determination of partial as well as full sequence information of the polynucleotide. That is, the terms include sequences of subsets of the full set of four natural nucleotides, A, C, G and T, such as, for example, a sequence of just A's and C's of a target polynucleotide. That is, the terms include the determination of the identities, ordering, and locations of one, two, three or all of the four types of nucleotides within a target polynucleotide. In some embodiments, the terms include the determination of the identities, ordering, and locations of two, three or all of the four types of nucleotides within a target polynucleotide. In some embodiments sequence determination may be accomplished by identifying the ordering and locations of a single type of nucleotide, e.g. cytosines, within the target polynucleotide "catcgc . . . " so that its sequence is represented as a binary code, e.g. "100101 . . . " representing "c-(not c)(not c)c-(not c)-c . . . " and the

like. In some embodiments, the terms may also include subsequences of a target polynucleotide that serve as a fingerprint for the target polynucleotide; that is, subsequences that uniquely identify a target polynucleotide, or a class of target polynucleotides, within a set of polynucleotides, e.g. all different RNA sequences expressed by a cell.

[0098] This disclosure is not intended to be limited to the scope of the particular forms set forth, but is intended to cover alternatives, modifications, and equivalents of the variations described herein. Further, the scope of the disclosure fully encompasses other variations that may become obvious to those skilled in the art in view of this disclosure. The scope of the present invention is limited only by the appended claims.

CLAIMS

What is claimed is:

1. A method of analyzing a polymer comprising:

translocating a polymer through a nanopore, wherein different kinds of monomers of the polymer are labeled with different optical labels that generate distinguishable optical signals and wherein the nanopore constrains the monomers to move single file through an excitation zone that encompasses a plurality of monomers;

detecting a time-ordered set of optical signals from the monomers as they pass through the excitation zone;

separating optical signals from different kinds of monomers to form monomer-specific time-ordered sets of optical signals; and

determining a sequence of monomers from the monomer-specific time-ordered sets of optical signals from the polymer.

2. The method of claim 1 wherein said polymer is a polynucleotide labeled by extending a primer annealed to a template nucleic acid molecule in the presence of labeled nucleoside triphosphates.

3. The method of claim 2 wherein said primer contains a key sequence with labeled nucleotides which generate an initial optical signal as said polynucleotide translocates said nanopore and passes through said excitation zone.

4. The method of claim 2 wherein said labeled nucleoside triphosphates comprise at least two distinguishable fluorescent labels attached to at least two different kinds of nucleoside triphosphates so that at least two different kinds of nucleotide in said extended primer may be identified by fluorescent signals generated by the distinguishable fluorescent labels.

5. The method of claim 1 wherein said polymer is a polynucleotide and wherein said step of determining includes forming candidate sequences from overlapping segments of nucleotides determined from said optical signals.

6. The method of claim 1 wherein said nanopore is a protein nanopore.

7. The method of claim 1 wherein said step of detecting further comprises exciting said optical labels in said excitation zone to generate said optical signals.

8. The method of claim 7 wherein said optical labels are fluorescent labels and wherein said excitation zone has a volume and geometry which are determined by said nanopore, mutual quenching of adjacent fluorescent labels and/or quenching agents.

9. A method of analyzing a polynucleotide comprising:

translocating a polynucleotide through a nanopore, nucleotides of the polynucleotide being labeled with fluorescent labels and the nanopore having a bore that spatially constrains the fluorescent labels to prevent emission of fluorescent signals during translocation thereof;

exciting the fluorescent labels;

detecting a time series of fluorescent signals from the fluorescent labels as the polynucleotide translocates through the bore; and

determining a sequence of fluorescent labels attached to nucleotides of the polynucleotide from the time series of fluorescent signals.

10. The method of claim 9 wherein said fluorescent labels on different kinds of nucleotides of said polynucleotide emit distinct fluorescent signals.

11. The method of claim 10 wherein each value in said time series of fluorescent signals comprises fluorescent signals from a plurality of said fluorescent labels.

12. The method of claim 11 wherein said step of detecting includes separating said distinct fluorescent signals to form a plurality of measured label-specific time series of fluorescent signals.

13. The method of claim 12 wherein said step of determining includes comparing said plurality of label-specific time series of fluorescent signals with nucleotide sequences and selecting a nucleotide sequence that would generate time series of fluorescent signals closest to said measured label-specific time series of fluorescent signals.

14. The method of claim 9 wherein said step of determining includes forming candidate sequences from overlapping segments of nucleotides determined from said fluorescent signals.

15. A method of analyzing a polynucleotide comprising:

translocating a polynucleotide through a nanopore, nucleotides of the polynucleotide being labeled with fluorescent labels and the nanopore having a bore with an entrance and exit, bore spatially constraining the fluorescent labels to prevent emission of fluorescent signals during translocation thereof;

exciting the fluorescent labels by an evanescent field that encompasses an entrance region adjacent to the entrance of the bore and an exit region adjacent to the exit of the bore;

detecting a time series of fluorescent signals from the fluorescent labels as the polynucleotide translocates the bore, the detected fluorescent signals comprising fluorescent signals of fluorescent labels in the entrance region and the exit region; and

determining a sequence of fluorescent labels attached to nucleotides of the polynucleotide from the time series of fluorescent signals.

16. The method of claim 15 wherein said fluorescent labels attached to different kinds of nucleotides generate distinguishable fluorescent signals.

17. The method of claim 16 wherein said step of detecting includes separating said distinguishable fluorescent signals and wherein said step of determining includes determining said sequence of fluorescent labels from time series of the separated fluorescent signals.

18. A method for determining a nucleotide sequence of a polynucleotide comprising the steps of:

translocating a polynucleotide through a bore of a nanopore, wherein nucleotides of the polynucleotide are labeled with fluorescent labels such that in free solution fluorescent labels of nucleotides are substantially quenched and wherein fluorescent labels within the bore are constrained such that substantially no detectable fluorescent signal is generated therein and wherein different kinds of nucleotide are labeled with different fluorescent labels that generate distinguishable fluorescent signals;

exciting the fluorescent label of each nucleotide upon exiting the nanopore and prior to quenching by interaction with a preceding mutually quenching fluorescent label or a quenching agent;

measuring fluorescent signals generated by fluorescent labels exiting the nanopore; and

separating fluorescent signals from different kinds of nucleotide to form nucleotide-specific time-ordered sets of fluorescent signals; and

determining a sequence of nucleotides from the nucleotide-specific time-ordered sets of fluorescent signals from the polynucleotide.

Epi-Illumination System

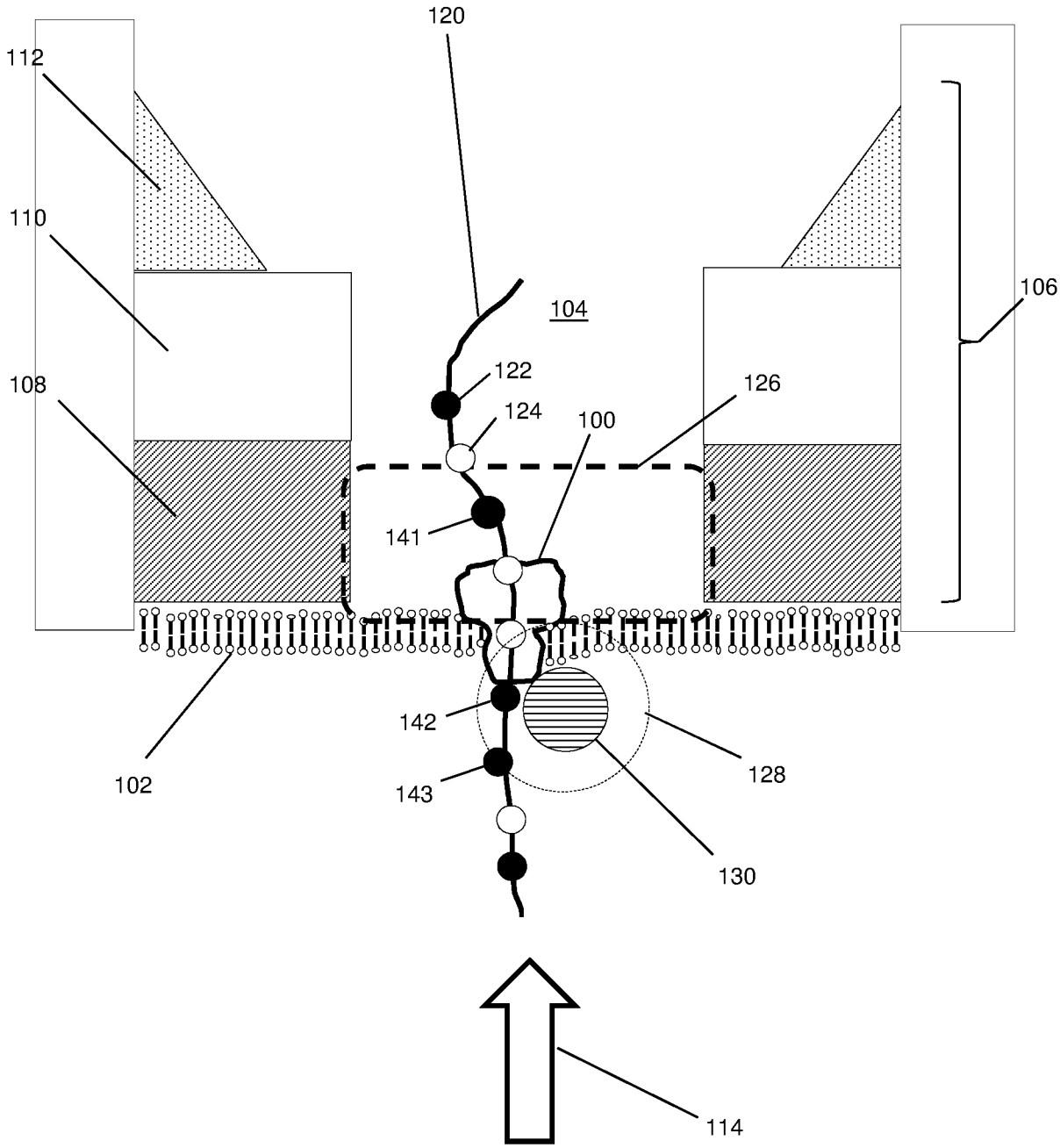


Fig. 1

TIRF System

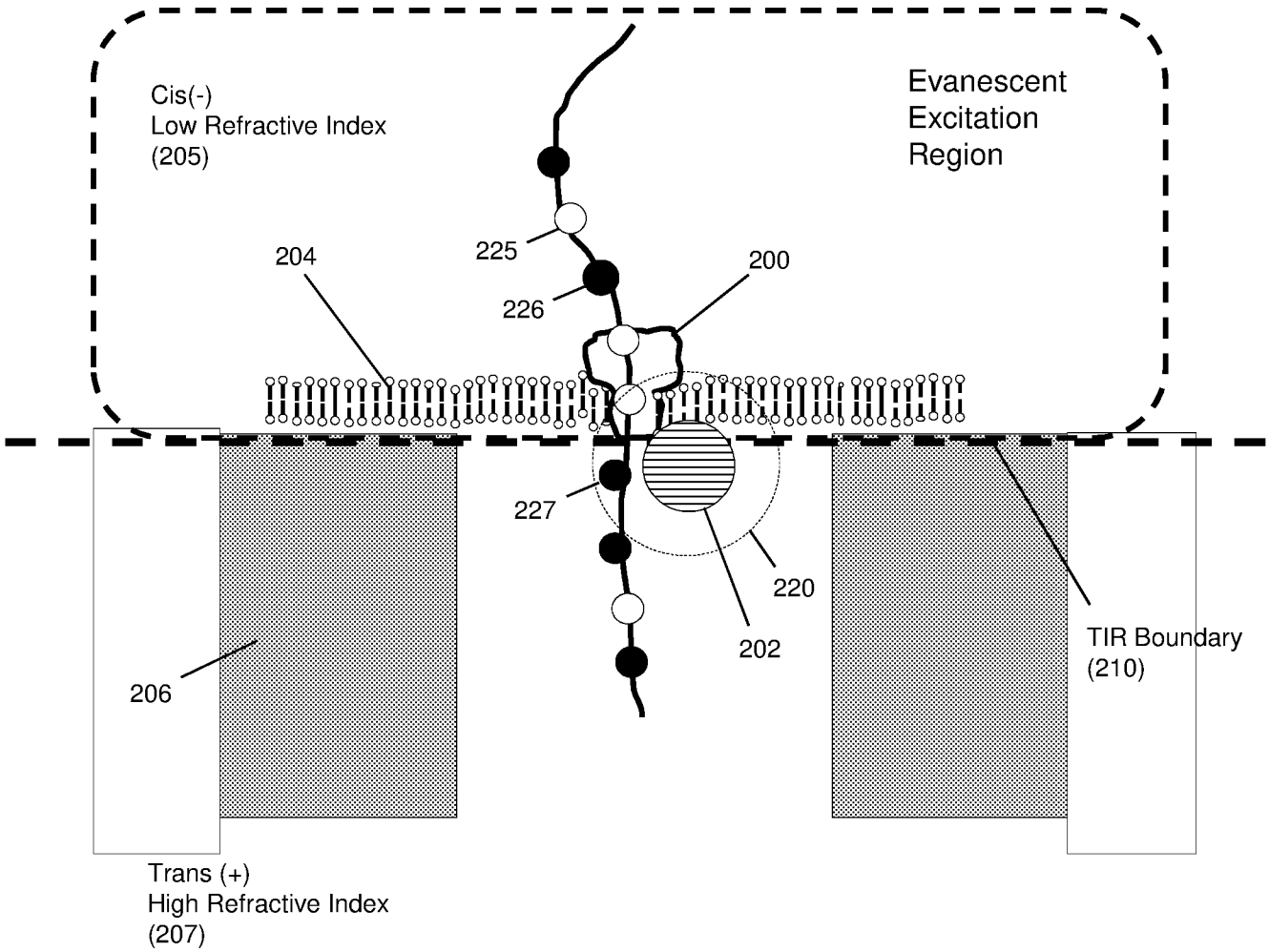


Fig. 2A

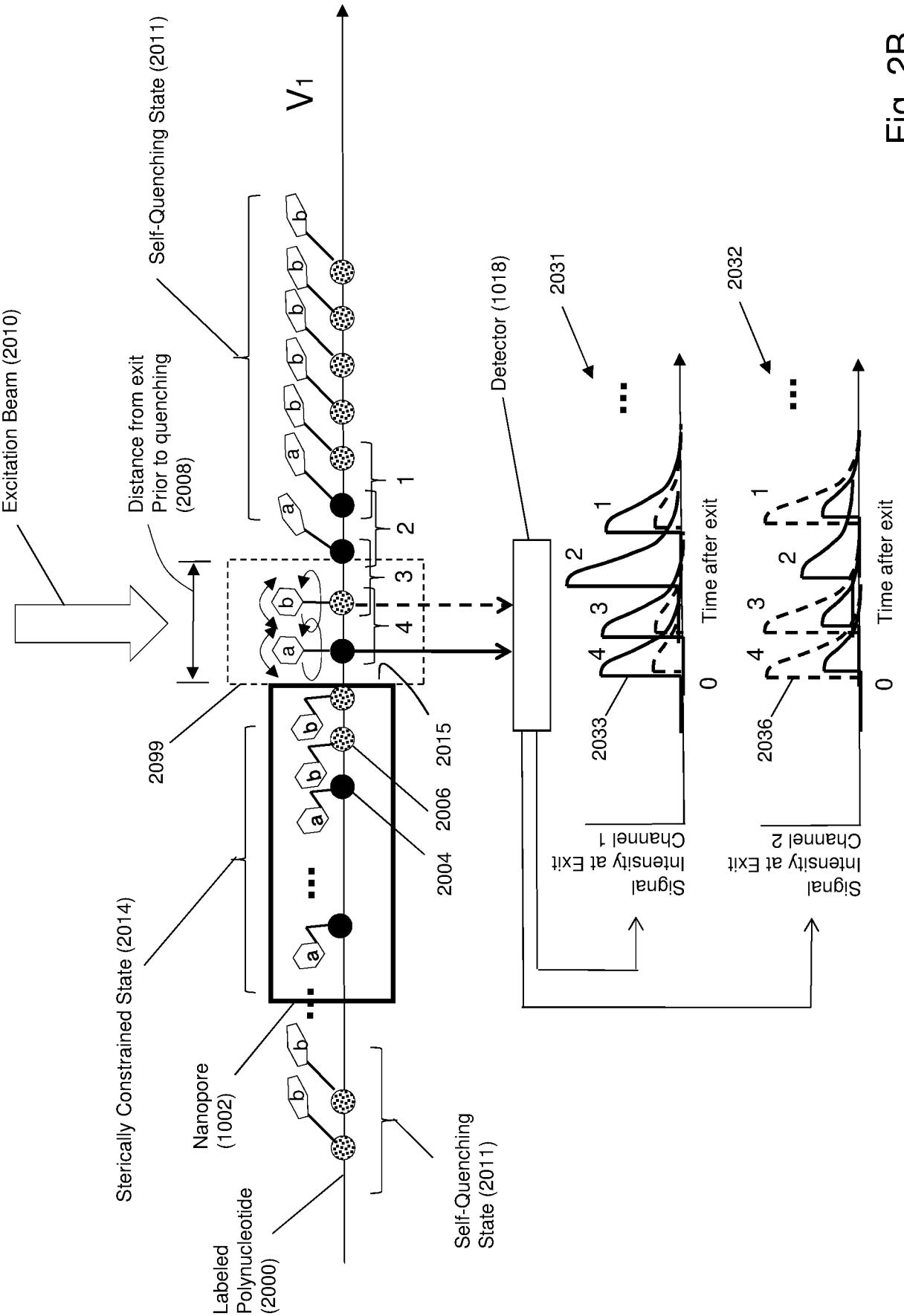

Excitation Beam (2010)
Distance from exit Prior to quenching (2008)
Self-Quenching State (2011)
V_1
Detector (1018)
2031
2032
Time after exit
Signal Intensity at Exit Channel 1
Signal Intensity at Exit Channel 2
2033
2036
2099
2015
2006
2004
Sterically Constrained State (2014)
Nanopore (1002)
Labeled Polynucleotide (2000)
Self-Quenching State (2011)

Fig. 2B

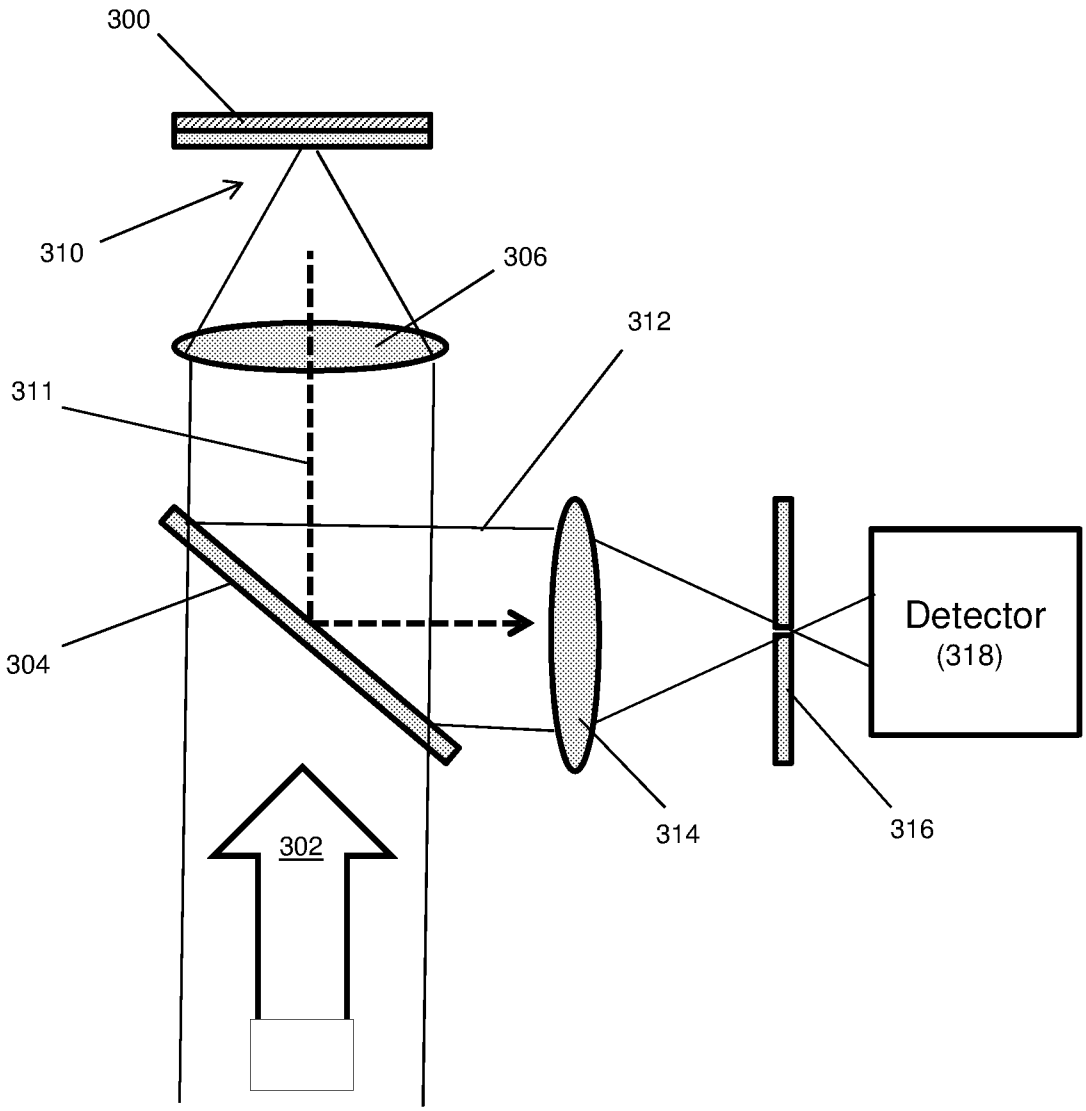


Fig. 3

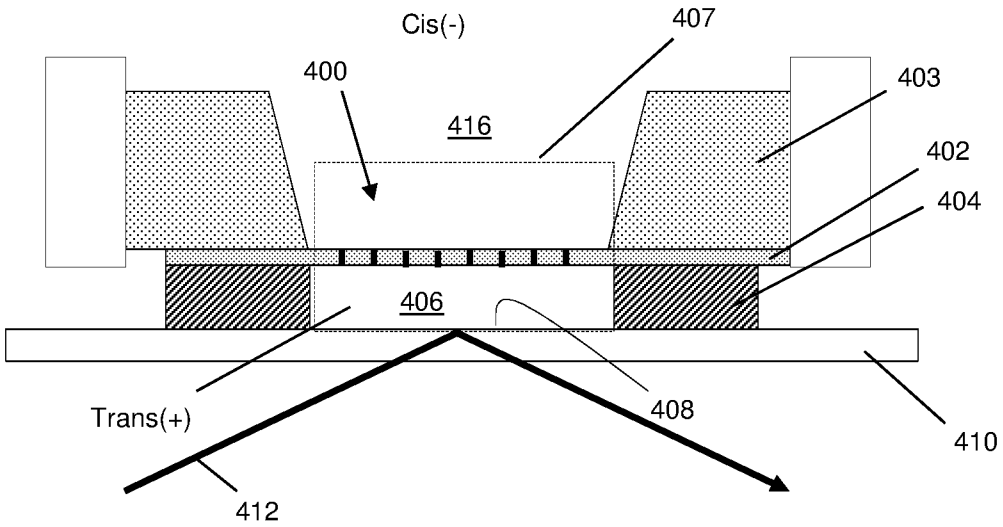


Fig. 4A

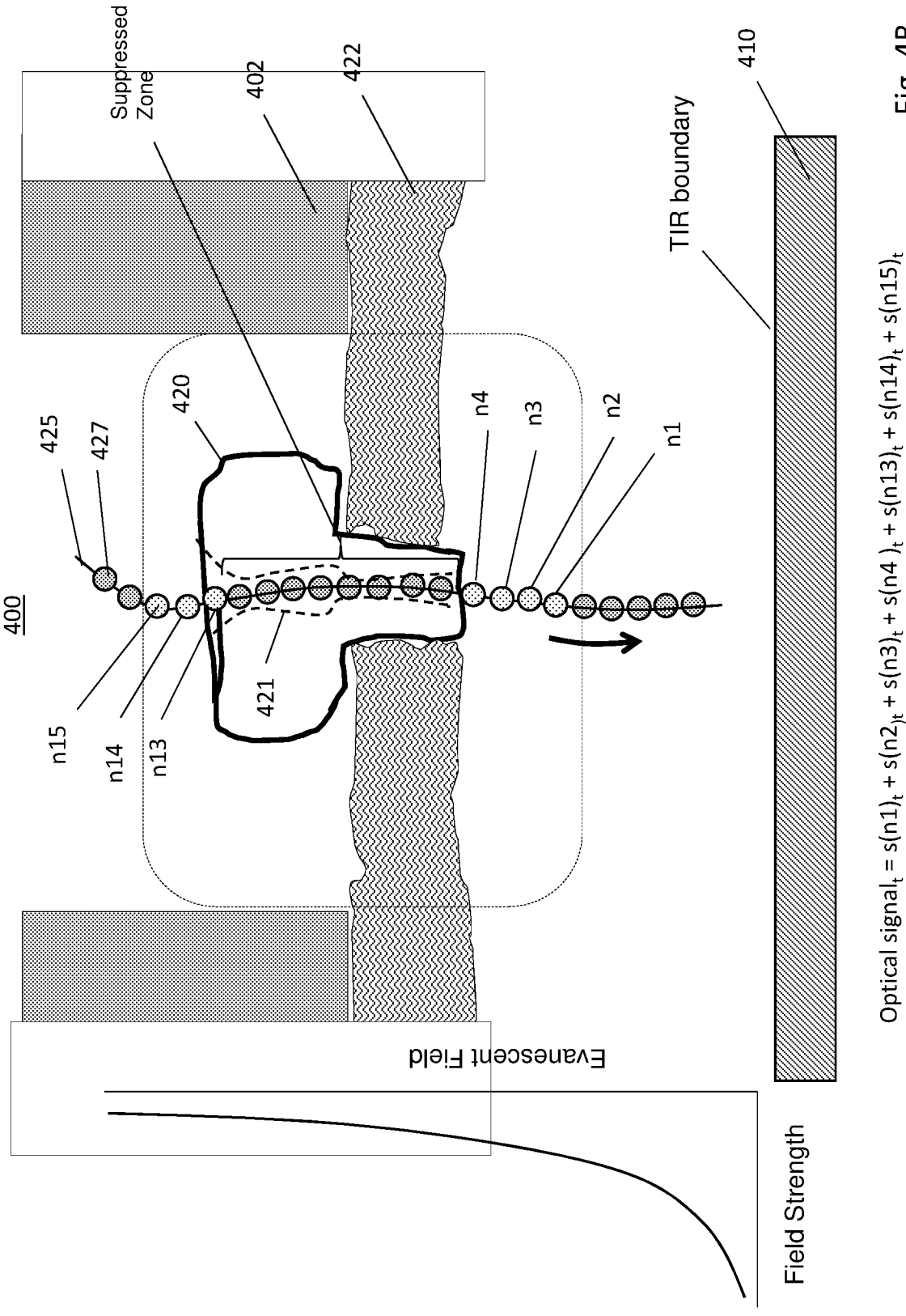

Suppressed Zone
402
422
410
TIR boundary
400
425
427
420
n4
n3
n2
n1
n15
n14
n13
421
Evanescent Field
Field Strength
Optical signal$_t$ = s(n1)$_t$ + s(n2)$_{jt}$ + s(n3)$_t$ + s(n4)$_t$ + s(n13)$_t$ + s(n14)$_t$ + s(n15)$_t$

Fig. 4B

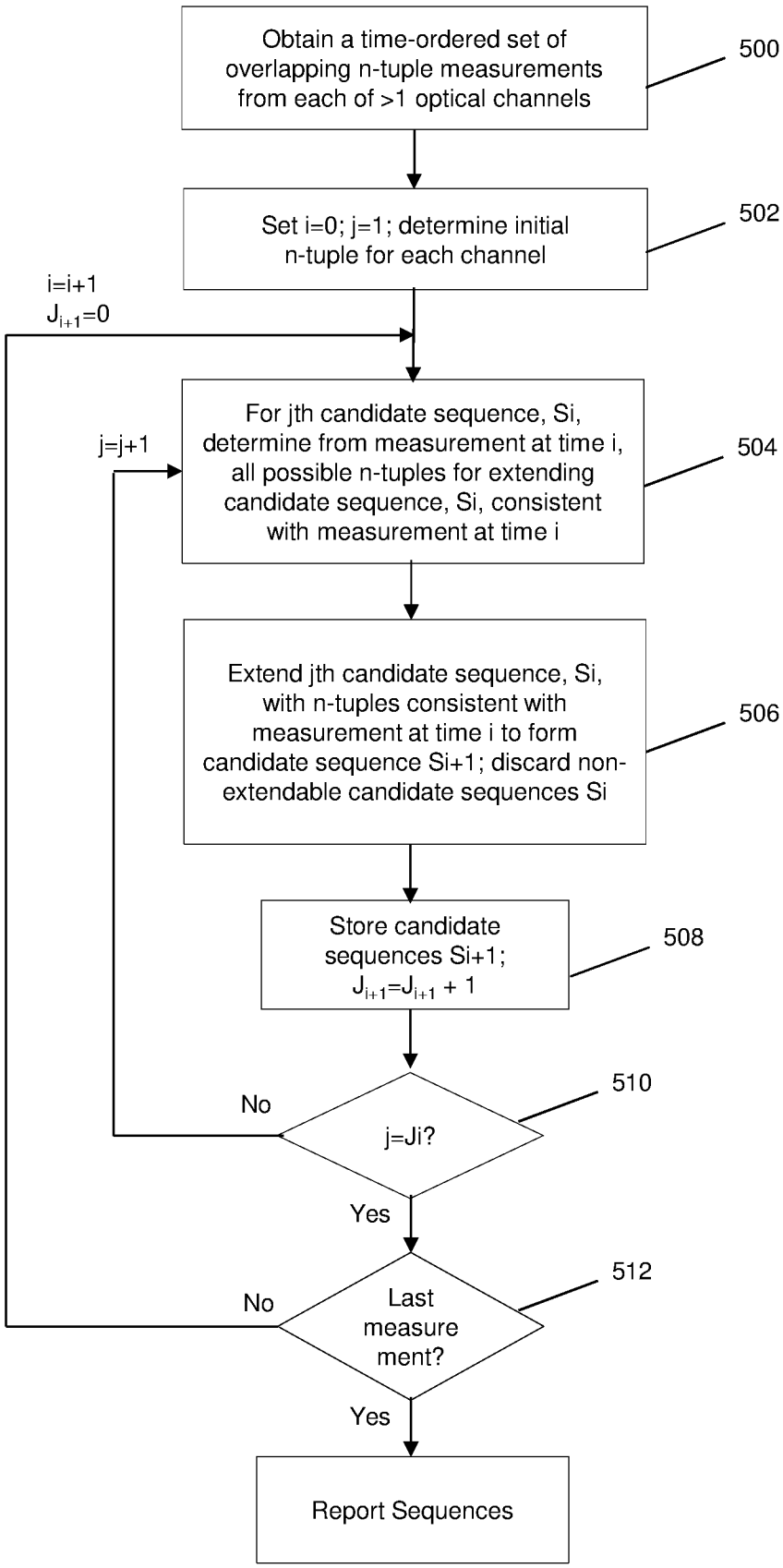

Obtain a time-ordered set of overlapping n-tuple measurements from each of >1 optical channels
500
Set i=0; j=1; determine initial n-tuple for each channel
502
i=i+1
J_{i+1}=0
For jth candidate sequence, Si, determine from measurement at time i, all possible n-tuples for extending candidate sequence, Si, consistent with measurement at time i
504
j=j+1
Extend jth candidate sequence, Si, with n-tuples consistent with measurement at time i to form candidate sequence Si+1; discard non-extendable candidate sequences Si
506
Store candidate sequences Si+1; J_{i+1}=J_{i+1} + 1
508
j=Ji?
510
No
Yes
Last measure ment?
512
No
Yes
Report Sequences

Fig. 5

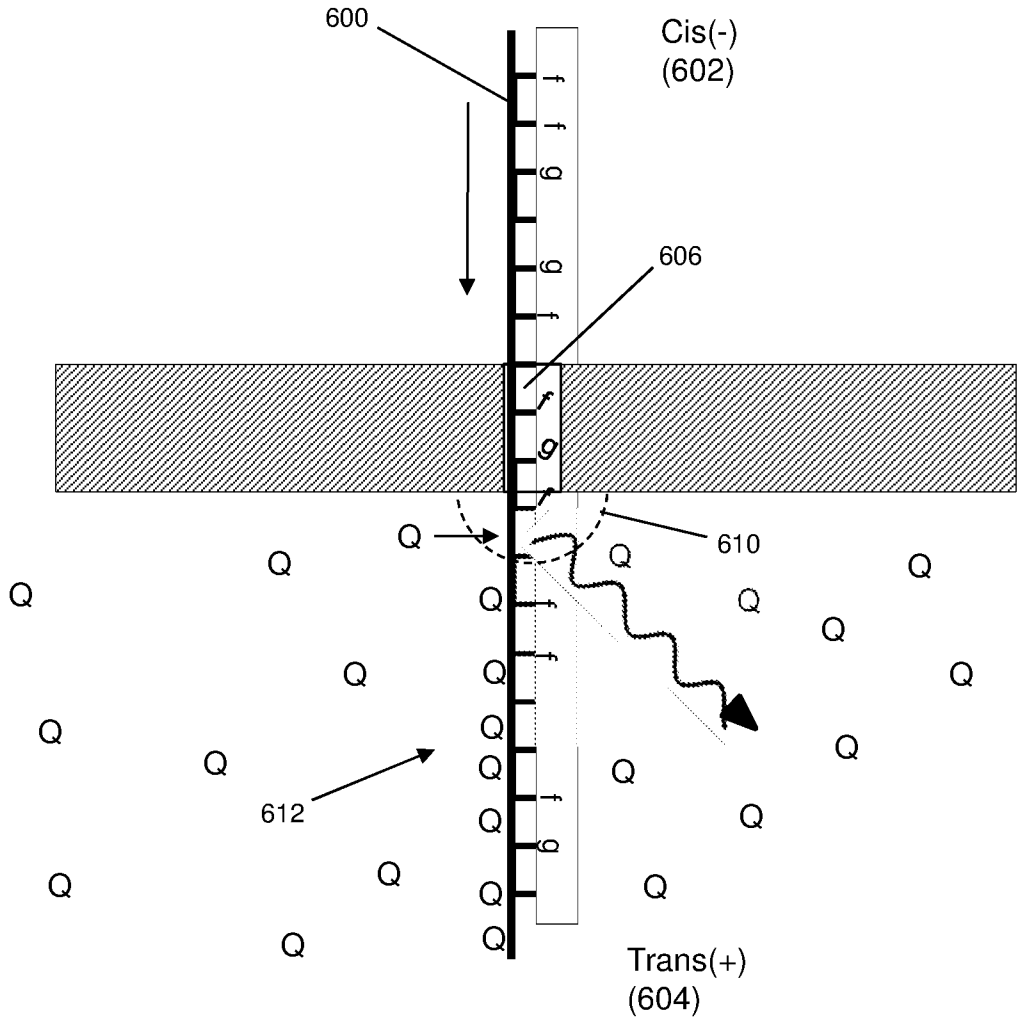
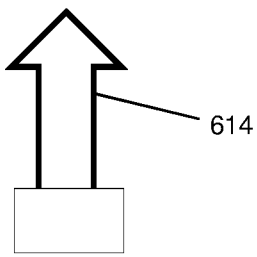


Fig. 6A



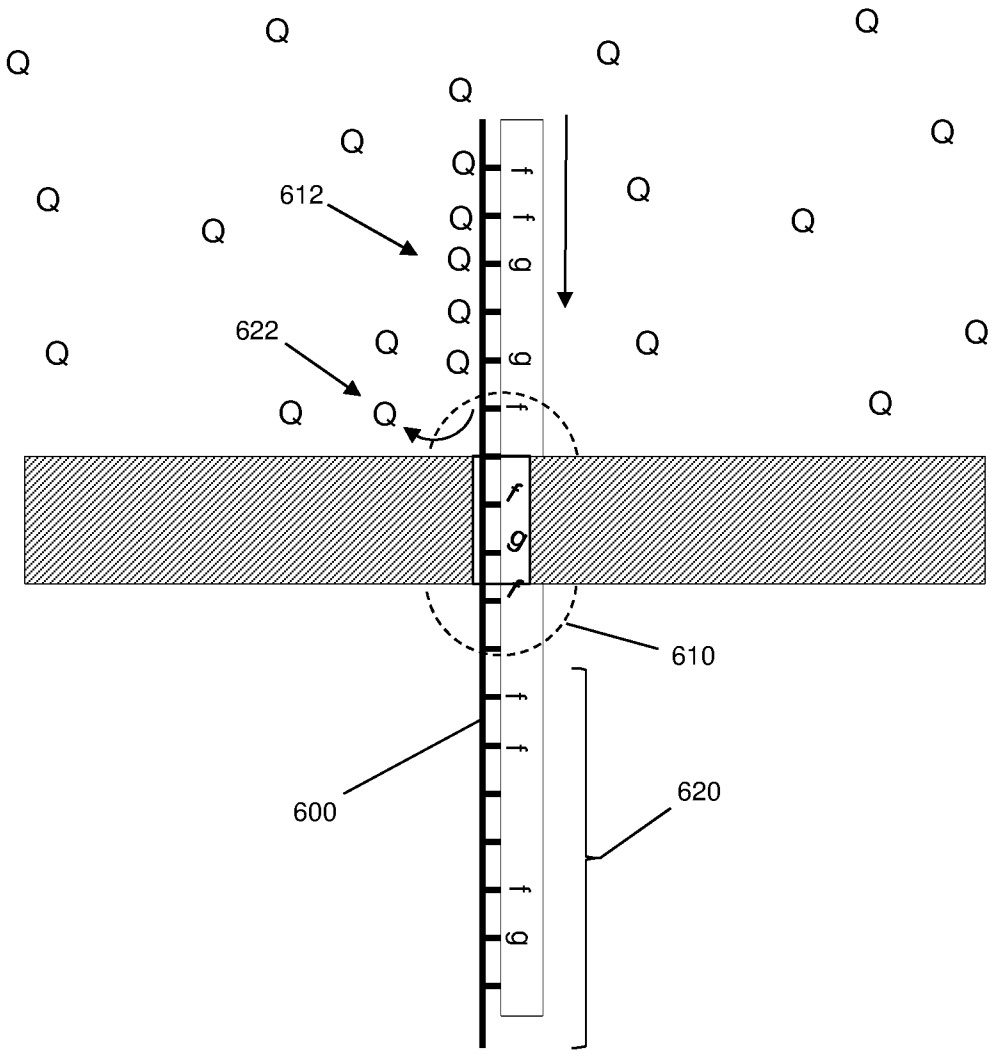


Fig. 6B

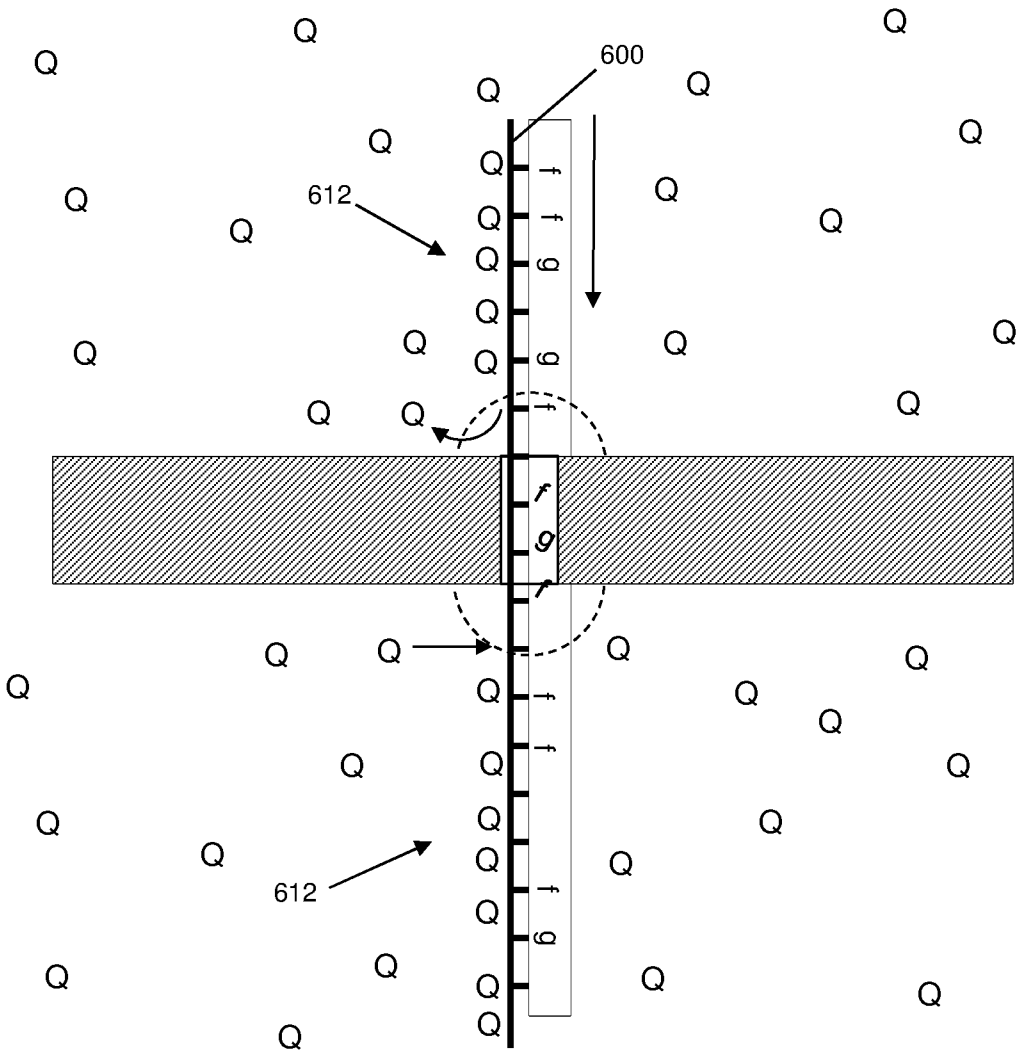


Fig. 6C

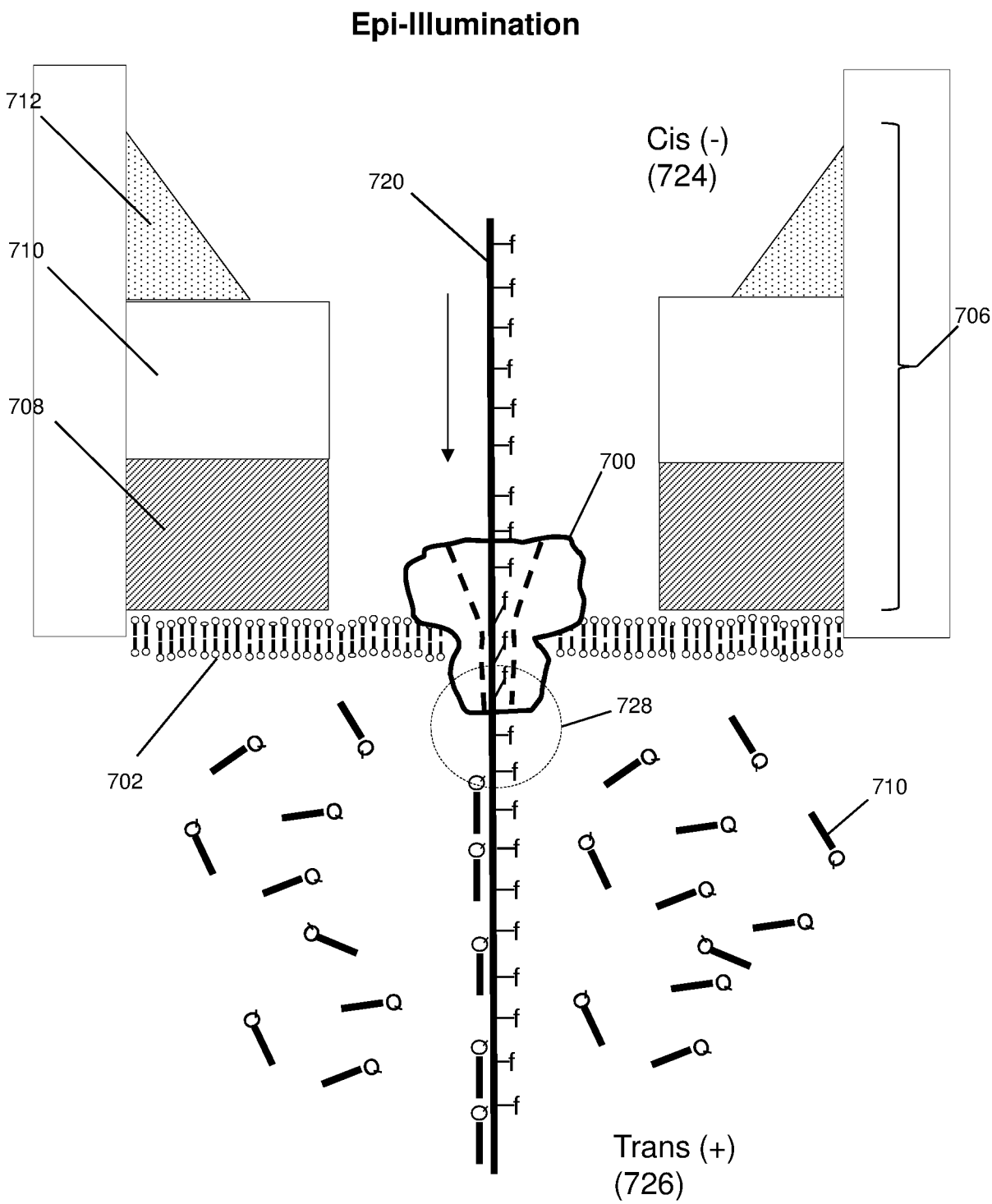

Epi-Illumination
712
710
708
720
Cis (-)
(724)
706
700
728
702
710
Trans (+)
(726)

Fig. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2017/024314

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12Q 1/68; C40B 20/00 (2017.01)
CPC - C12Q 1/6869; C40B 20/00; G01N 21/6428 (2017.02)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC - 506/2; 506/4; 977/774 (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 9,279,153 B2 (HUBER) 08 March 2016 (08.03.2016) entire document	1, 5-8
---		---
Y		2-4, 18
X	WO 2014/190322 A2 (QUANTAPORE INC) 27 November 2014 (27.11.2014) entire document	9-12, 14-17
---		---
Y		13, 18
Y	US 7,575,865 B2 (LEAMON et al) 18 August 2009 (18.08.2009) entire document	2-4
Y	US 2013/0040827 A1 (MACEVICZ) 14 February 2013 (14.02.2013) entire document	4, 13

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 May 2017	22 JUN 2017
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300	Authorized officer Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774